Orange County Saturation Study

A Study by the

Metropolitan Water District of Southern California

and the

Municipal Water District of Orange County

July 24, 2002

Table of Content

List of Appendices

List of Tables

PREFACE

The *Orange County Saturation Study* was jointly funded by the Metropolitan Water District of Southern California and the Municipal Water District of Orange County. A grant from the United States Bureau of Reclamation assisted the overall effort. The help of Ms. Meena Westford in obtaining this grant is greatly appreciated.

Several individuals contributed to this study. Dr. Michael Hollis of Metropolitan formulated the basic research design, drafted early versions of the survey instruments, and coordinated data analysis and report writing. Dr. Anil Bamezai of Western Policy Research served as project manager. Dr. Bamezai assisted in developing the research design; monitored all data collection activities; performed the statistical analysis used in estimating showerhead saturation and program freerider rates; and drafted sections of the final report. Dr. Charles DiSogra of the University of California at Los Angeles reviewed the telephone survey instrument and provided valuable advice on sample design and weighting issues. Mr. John Wiedmann and Mr. Warren Teitz of Metropolitan helped develop the residential inspection survey instrument. In addition, Mr. Wiedmann worked on the protocol used for the residential on-site inspections and helped resolve several field and logistical issues.

Farrand Research of Long Beach, California, administered the telephone screening survey. Mr. Tom Farrand, Dr. Mary Barton and Mr. Eric Nelson helped draft the questionnaire and administered the survey. Peak International of Yorba Linda, California, scheduled and conducted the on-site inspections.

The *Orange County Saturation Study* would not have been possible without the contributions of Mr. Joe Berg, Water Efficiency Programs Manager at the Municipal Water District of Orange County (MWDOC). Mr. Berg served as the liaison between the study team and Orange County retail water agencies, effectively attending to any concerns raised by these agencies or their customers. Mr. Steve Hedges of MWDOC assisted in this effort and also helped resolve several field issues. Finally, Ms. Pat Meszaros of MWDOC fielded calls from several recalcitrant study participants. The information and reassurance she provided was instrumental in persuading these households to participate in the study.

EXECUTIVE SUMMARY

The *Orange County Saturation Study* had two main objectives. The first was to test the hypotheses that: (1) the replacement of water-wasting showerheads with low-flow models has occurred in a geographically uniform pattern among Orange County households; and (2) that 75% of all residential showerheads are low-flow devices. The second objective was to develop a representative profile of the types of water using devices found in the typical Orange County home.

The conservation activities of urban water agencies in California are largely shaped by the California Urban Water Conservation Council's (CUWCC) 14 urban conservation "best management practices" (BMPs). One of these requires the distribution of low-flow showerheads (LFSHs) and other water-conserving indoor plumbing devices.[1] One means available to utilities for demonstrating that they have fully satisfied the requirements of this BMP is to show, using survey data, that 75% of all residential showerheads in a retail agency's service territory have been fitted with low-flow showerheads.

A principal objective of the *Orange County Saturation Study* was to test the hypothesis that, at least within Southern California, the results from countywide surveys will provide good estimates of the saturation of LFSHs among households served by the individual retail agencies comprising the county. The study results strongly support this hypothesis but suggest that the required 75% saturation threshold had not quite been achieved as of the winter of 2000. The estimated saturation rates were, however, close enough to the requisite 75% level to suspect that the combined effect of utility distribution programs and natural replacement (i.e., the replacement of old or broken devices at the occupant's or owner's expense) may have substantially bridged the gap between then and now.

The second major study objective was to develop a representative profile of the incidence and use of other indoor water-using devices in pre-1992 units.[2] The survey revealed important differences in the saturation of low-flow showerheads and ultra-low flush toilets (ULFTs) in single-family and multi-family dwellings. It was also found that the reasons for replacing old water-wasting devices with their more efficient alternatives tend to differ in important respects between the two types of households. Combined, these differences suggest useful lessons for the future design and targeting of utility-funded residential plumbing retrofit programs.

[1] See California urban Water Conservation Council, *Memorandum of Understanding Regarding Urban Water Conservation in California*, 18-20. Note that ultra-low flow toilet retrofits are covered by a separate BMP.
[2] The focus was on the pre-1992 housing stock because state law has, since 1992, prohibited the sale or installation of non-conserving showerheads, toilets and faucet aerators.

I. INTRODUCTION

Water agencies in Southern California have aggressively funded a variety of conservation programs for more than a decade. During this time, for example, the Metropolitan Water District of Southern California (Metropolitan) and its member agencies have spent more than $240 million on residential indoor water use efficiency programs. Among other achievements, these programs have resulted in the distribution of more than 3 million low-flow showerheads (LFSH) and the retrofit of approximately 2 million water-wasting toilets with ultra-low flush (ULFT) models.

Given this considerable investment, an obvious question arises about the current saturation of water-efficient showerheads and toilets among Southern California households. The answer has important implications for the cost-effectiveness of current utility-funded residential retrofit programs.

Since 1992, California law has prohibited the sale and installation of non-conserving showerheads and toilets.[3] Accordingly, all water-wasting showerheads and toilets will eventually be replaced with water-efficient alternatives, at private expense, as these devices fail with age and as bathrooms are remodeled. Within this context, the primary objective of utility-funded plumbing retrofit programs is to accelerate the rate at which water-wasting devices are replaced with water conserving alternatives relative to what would occur under the process of *natural replacement* alone. As long as the value of these accelerated water savings exceed the costs of utility-funded retrofit programs, it is economically beneficial—from the utility's perspective—to fund these types of programs.[4] Once some threshold level of utility-assisted device saturation has been achieved, however, it becomes prudent to rely on natural replacement to change-out the remaining stock of water-wasting devices, allowing utility funds to be reallocated to programs capable of producing greater returns per conservation dollar invested. The existing absence of reliable information about the current saturation of low-flow showerheads and ultra-low flush toilets is, however, a significant impediment to ensuring that the current portfolio of utility conservation investments is yielding optimal returns.

The current study had to primary objectives. The first was to test the hypothesis that households served by utilities in Orange County, California, have achieved the 75% low-flow showerhead saturation level required to document full implementation of the California Urban Water Council's (CUWCC) residential plumbing retrofit conservation *best management practice*

[3] Technically, this law allowed restrictions on the installation of non-conserving toilets to be phased in over a two-year period. This provision was intended to ease the economic hardship on retailers and developers with large inventories of non-conserving toilets. For purposes of this study, however, it is assumed the law took effect in 1992.
[4] It is currently estimated that replacing a non-conserving toilet with a ULFT in a typical single-family home in Southern California saves about 24 gallons of water per day. Accordingly, a utility-funded ULFT retrofit program that accelerates the replacement of a single non-conserving toilet by five years, for example, would save approximately 43,800 gallons of water during this five-year period.

(BMP). The second objective was to develop a representative profile of the types of indoor water using devices found in the typical Orange County residence.[5] These two objectives are discussed at greater length below.

A. Study Objectives

1. Low-Flow Showerhead Saturation Hypothesis

Under the CUWCC's current *Memorandum of Understanding Regarding Urban Water Conservation* (*MOU*), signatory agencies can be certified as having met the implementation requirements of urban conservation BMP 2 by demonstrating

> through customer surveys with 95% statistical confidence and a ±10% error that 75% of single-family residences and 75% of multi-family units constructed prior to 1992 are fitted with low-flow showerheads. [6]

This provision was adopted during the 1997 urban *MOU* revision process. Shortly thereafter, Metropolitan asked that Southern California be allowed to use the results from county-level sample surveys to estimate the level of low-flow showerhead (LFSH) saturation among households served by the separate retail water agencies within the county.

Metropolitan's request was motivated by the belief that the distribution and installation of LFSHs has been spatially uniform throughout the urban part of Southern California. There are two main reasons for this belief. First, Metropolitan and its member agencies have distributed over three million LFSHs in Southern California since 1990. While some of these have been directly installed as part of utility-funded residential survey programs, a sizable share of these devices were given away during the early 1990s via regional mass distribution programs. These programs were administered not only by water agencies, but also by regional gas and electric utilities. Schools, fast food outlets, building supply stores and special promotional events throughout Southern California were also used as major distribution channels. As such, many LFSH distribution programs were regional in scope.

A second reason for expecting that residential LFSH retrofits have occurred uniformly throughout Southern California involves the previously mentioned 1992 state law restricting the manufacture, sale and installation of showerheads to low-flow models. This means that occupants have no option other than replacing old or broken water-wasting showerheads with a water-efficient model. Given the relatively short physical device life of most showerheads, the process of natural replacement contributes to the saturation of LFSHs over time and that the

[5] Orange County was selected as the site for testing the research design used by the study. As will be discussed, the study results suggest that the basic methodology is transferable to other geographic areas.
[6] California Urban Water Conservation Council, *Memorandum of Understanding Regarding Urban Water Conservation in California,* Exhibit 1, p. 21.

incidence of these retrofits should be generally similar among the individual cities and towns comprising a county.

If the geographic uniformity hypothesis is correct, a random sample of households drawn from one part of a county should produce saturation estimates similar to what would be obtained from a random sample of households drawn from any other part of the county. Likewise, a well-designed countywide sample of households should produce saturation estimates that are statistically equivalent to those obtained from a sample drawn from any city or similar geographic sub-unit within the county. The implications of this conjecture are important. While the retail water agencies operating within a county might be able to collectively finance a countywide sample, only the largest of urban agencies can afford to fund these types of surveys within their own service areas. Furthermore, the prospect of agencies diverting money from conservation programs to fund separate surveys to develop saturation estimates that might be suitably obtained from a single countywide survey is clearly an inefficient use of conservation resources.

For these reasons, Metropolitan asked CUWCC's Measurement and Evaluation Committee to be allowed to use the results from countywide surveys to estimate LFSH saturation among households served by the individual retail water utilities operating within each county served by Metropolitan. While the Committee acknowledged the potential merits of this request, it asked Metropolitan to provide empirical support for the geographic uniformity hypothesis. In response, Metropolitan proposed a method of operationally testing this hypothesis. The Committee subsequently approved the test protocol. The Committee also agreed to recommend that CUWCC's Steering Committee allow Metropolitan, or its member agencies, to use the results of countywide surveys as a basis for estimating saturation at the retail agency level if the test results supported the hypothesis of geographically uniform LFSH saturation rates.

2. Residential Indoor Plumbing Device and Usage Profile

Although the *Orange County Saturation Study* was initially designed to test the geographic uniformity hypothesis and to estimate the current saturation of low-flow showerheads, it was recognized that incidence and use of other indoor water using devices could be collected at relatively small incremental cost. Given of the value of this type of information to conservation planners, the study was expanded to include a second major objective—developing a representative profile of the types of indoor water using devices found in the typical Orange County home.

B. Report Overview

Chapter II discusses the basic research design used by the study. This includes the rationale underlying the test used to evaluate the geographic uniformity hypothesis. The sampling design, definition of measurement of key variables, sample size requirements, and data collection strategies and outcomes are also discussed.

- Chapter III considers the representativeness and quality of the study sample. The reliability of self-reported information obtained from surveyed households is also addressed.

- Chapter IV presents a formal test of the geographic uniformity hypothesis. As is shown, the results provide compelling support for the hypothesis—at least within Orange County. Separate estimates of the current saturation of low-flow showerheads are also developed for single- and multi-family dwellings in Orange County.

- Chapter V profiles the types of indoor plumbing devices found in the typical Orange County home. Notable differences between single-and multi-family units are highlighted and their possible implications for conservation program and outreach efforts are discussed.

- Chapter VI summarizes the most important findings, provides necessary qualifications, and suggests ways for improving similar studies in the future.

Readers who are mainly interested in study findings can skip to Chapter IV. Those who are interested in the research design and data collection methods should read Chapters II and III and also consult the appendices.

II. STUDY DESIGN AND IMPLEMENTATION

The effective design and targeting of utility-sponsored residential plumbing retrofit programs requires accurate information about the existing type and distribution of household water-using devices. Similarly, the accuracy of water demand forecasts developed from end-use models depends critically on the availability and quality of information about the saturation of water-efficient devices, and about the rate at which homeowners replace water-wasting devices with their more efficient alternatives.

Most existing estimates of the incidence of different types of household water-using devices are based on information collected during the course of utility-sponsored residential surveys, or audits. But the ability to meaningfully generalize estimates based on households who voluntarily participate in these programs is often limited by the fact that program participants do not constitute a representative sample of the utility's residential customer base.

This chapter describes the basic methodology and procedures used in the study. Primary attention is given to the conceptual, methodological and logistical issues involved in identifying, recruiting, and collecting information from a representative sample of households located in an area served by multiple retail water agencies. First, however, design considerations specific to the task of testing the hypothesis that the saturation rate of low-flow showerheads (LFSH) is geographically uniform throughout Orange County will be discussed.

A. Formulating the Low-Flow Showerhead Saturation Research Hypothesis

At the request of CUWCC's Steering Committee, Metropolitan developed the following three alternatives for testing the geographic uniformity hypothesis.

1. Randomly select up to four retail agencies within a single county. Conduct separate surveys for each of these agencies and for the county as a whole. Test to determine if saturation estimates obtained from the individual agencies and the county surveys jointly satisfy the 75% saturation requirement.

2. Randomly select two retail agencies within two randomly selected counties. Conduct separate surveys for each of the agencies and each county. Conduct the same tests as would be used under option (1).

3. Randomly select two counties. Within each county, purposely select two retail agencies with historically low levels of agency-sponsored conservation program activity. Conduct the same tests as would be used under option (1).

The matter was referred to CUWCC's Measurement and Evaluation Committee, which expressed a strong preference for the first alternative. Metropolitan then suggested a variation on

this design. Specifically, it proposed that purposive selection of two cities within the county having relatively low levels of utility-funded conservation activity within the county would provide a stronger test of the uniformity hypothesis. This is because the rate of showerhead retrofits within communities served by these types of utilities should approximate what would occur under conditions of natural replacement combined with a regional distribution program. It was also proposed that, to minimize costs, the study consider only two retail agencies. The Measurement and Evaluation Committee agreed to the revised proposal and forwarded its recommendation to the Steering Committee. It asked Metropolitan to develop an appropriate research design. Metropolitan also assumed responsibility for decisions about data collection methods and for developing suitable operational definitions and measures of key concepts and variables.[7]

B. Study Site Selection

Metropolitan's service area is composed of all or parts of Los Angeles, Orange, San Diego, Ventura, San Bernardino and Riverside counties. Orange County was selected as the study site for several reasons. First, the Municipal Water District of Orange County's Water Use Efficiency Programs Manager was willing to serve as a liaison between the research project and the retail water agencies operating within the county. Second, the county has several cities that are served by their own retail water utilities. This greatly simplified the logistics and the cost of data collection without compromising the basic integrity of the study design.[8] Among the possible candidates, the cities of Buena Park and Fullerton were ultimately selected. Among other considerations, these two cities satisfied the study's needs for sites having relatively low historic levels of local utility-funded showerhead distribution programs, relatively large populations, and a desirable mix of pre-1992 single- and multi-family housing units.

[7] It is worth noting that the language in the urban *MOU* describing the LFSH coverage requirement is somewhat vague. According to the *MOU*, the criteria for demonstrating BMP 2 compliance using survey data requires showing that "75% of single-family residences and 75% of multi-family units constructed prior to 1992 are fitted with low-flow showerheads" (CUWCC *MOU*, p. 20). This language could arguably be interpreted as meaning that as long as a unit has *at least one* low-flow showerhead, it satisfies the requirement that the *unit* be "fitted with low-flow showerheads." This standard is easier to meet than the one that results if the language is interpreted as requiring that 75% of *all* showerheads in a dwelling be fitted with low-flow devices. For purposes of this study, it is assumed that the *MOU* language was intended to impose the more stringent coverage requirement. As such, the basic unit of analysis for estimating and testing LFSH saturation rates will be showerheads, not housing units.

[8] As discussed in Section II.E, a telephone screening survey was conducted to identify eligible study households. For households comprising the retail agency samples, it was necessary that the household receive its water service from one of the two retail agencies participating in the study. Since apartment dwellers often do not receive or directly pay their water bill, it is not uncommon for individuals residing in these types of units to not know the name of the retail agency that provides their water. On the other hand, most full-time household residents can accurately identify the city in which their dwelling is located. The overlap between city and retail water agency service boundaries permits eligibility screening to be based on answers given to a question about city of residence rather than the name of the retail water supplier.

C. Definition and Measurement Issues

1. Defining and Distinguishing Single- and Multi-Family Dwellings

Although the urban *MOU* clearly specifies that separate saturation estimates are required for single- and multi-family units, the document does not explicitly define or operationally distinguish these two types of dwellings. It was therefore necessary to develop operational definitions suitable for reliably classifying dwellings for purposes of the study.

Most utilities distinguish between customers occupying single-family and multi-family dwellings. But they do so in different ways. The most common practice is to rely on the way the service is metered: separately metered units are classified as single-family dwellings and master-metered units are classified as multi-family dwellings. But there are variations. Some utilities use a combination of metering type and structural characteristics. For example, separately metered free-standing housing units and master-metered units consisting of four or fewer residential units might be classified as single-family dwellings, while all master-metered and separately metered dwellings consisting of five or more units would be classified as multi-family.

Whereas utilities tend to rely on meter-based definitions, the U.S. Census and the California Department of Finance (DOF) define dwelling types in terms of structural characteristics. Freestanding structures and residential dwellings sharing a common wall extending from the ground to the roof are classified as single-family structures. All multi-story buildings in which individual dwelling units have a separate residence located either above or below are classified as multi-family.

While the meter-based definition has some appeal, the DOF definition was used by the *Orange County Saturation Study*. There were several reasons for this. First, a meter-based definition would have required obtaining customer billing records from all retail utilities serving the county to use as a sampling frame.[9] That would have been time consuming and expensive, and the unwillingness of a single utility to provide this information would have resulted in an incomplete frame. Second, the DOF annually provides population and housing estimates for cities and counties. This information can be extremely useful in assessing how well a sample represents the existing population of housing units and for developing the sample weights that are often used to statistically correct for common imbalances between the sample and the parent populations. To use DOF data for this purpose, however, required using DOF definitions to classify sampled households.[10]

[9] The meaning and function of a sampling frame is discussed in a later section.

[10] Information about metering status was collected during the survey. While the results are discussed in this report, a preliminary analysis based on metering status did not produce substantively different conclusions from those discussed in Chapter III. The issue merits further investigation.

2. Measurement Issues

Good measures share two basic properties—*reliability* and *validity*.[11] Simply stated, a reliable measure is one that produces the same value on repeated applications under identical conditions.[12] Since few measures are perfectly reliable, the values obtained from a given measurement procedure will tend to differ somewhat. Generally speaking, however, the more reliable the method used to record measurements, the smaller the difference, on average, between the measured values and the corresponding *true* value of the item or attribute being measured.

Measurement procedures producing measures that, on average, systematically over-estimate or under-estimate the true underlying population value of a characteristic suffer from non-random measurement error. Because of these systematic differences, the corresponding measurement procedure is said to lack validity. In this sense, statistics calculated using sample data contaminated by non-random measurement error produce invalid, or biased, estimates of the population parameters of interest.

Consider, by way of example, the task of classifying a showerhead as either an ultra-low flow device or as a less water-efficient model. One strategy would be to base the classification on simple visual inspection. Alternatively, since LFSHs are defined as showerheads having flow rates of 2.5 gallons per minute (gpm) or less, a second possibility would be to directly measure the showerhead's flow rate and to base the classification on this measurement.

If LFSHs were labeled as such, or had distinct physical characteristics, visual inspection alone might yield highly reliable information about the saturation of low-flow showerheads. Unfortunately, since these devices are not labeled and do not have unique visual characteristics, classifications based on visual inspection alone would not normally be especially reliable or valid.

The point of this discussion has been to introduce some terminology that will be used later. It also serves to highlight the fact that careful thought needs to be given to both the methods and the measures used to collect data to be used in estimating and comparing saturation rates. Measures that are both unreliable and invalid will typically result in incorrect statistical test results and erroneous substantive interpretations.

[11] For a non-technical discussion of measurement reliability and validity with specific application to the evaluation of water conservation programs, see Hollis, Bamezai, and Pekelney (1998). For more detailed statistical treatments, see Carmines (1980) and Hollis (1991).

[12] Technically speaking, reliability is a property of the entire process used to record the value of some characteristic of interest. For measures obtained by an opinion survey, this would include the question wording and response options used to measure an individual's opinion about a given issue, the context in which the question is asked, and the method used to administer the survey (e.g., telephone, mail or in-person), among other factors. It is this entire process of measurement that is implied whenever the term "reliability" is used in this discussion.

3. Alternative Data Collection Methods

Two methods have commonly been used to generate the information needed to estimate the saturation and spatial distribution of water-efficient devices—telephone surveys and on-site inspections.[13] Each has distinct advantages and disadvantages.

Telephone Surveys

Telephone surveys have sometimes been used by utilities to estimate the saturation of water-using devices.[14] The advantages of this method include lower costs per surveyed household, fast turn-around, and the potential for greater quality control, since all calls are made from a single location, allowing closer supervision and quick problem resolution.

Serious questions exist, however, about the ability of a telephone survey alone to provide data of sufficient quality to yield good device saturation estimates. First, the survey research literature amply demonstrates that the answers people give to survey questions (sometimes referred to as *self-reports*) are often incorrect. While this is sometimes the result of intentional deception (e.g., respondents may either overstate or understate their education and income, depending on the context of the interview),[15] response errors are often the product of faulty memory or lack of knowledge. Second, it is often difficult for the typical person to reliably distinguish water-

[13] ***Data loggers*** represent a new and potentially promising data collection method. Because these devices can be unobtrusively attached to the customer's meter without the need to obtain the occupant's permission, the method may hold the greatest potential for collecting data from a truly random sample of household. Second, given an appropriate research design, a relatively small number of loggers could be used to gather information from a large sample of households. In this way, data logging could prove a far cheaper method of data collection than on-site inspections. Third, the use of loggers allows the possibility of monitoring changes in plumbing devices and water use behavior over time.
Despite these potential advantages, the use of data loggers also raises several questions. First, more research is needed to establish the reliability and validity of the end-use measures obtained from the logging devices and of the software used to decompose these measures into the discrete water using devices drawing water though the main service connection. Second, to the extent that the study objective is to develop predictive models that "explain" device replacement and water using behavior, additional information would need to be collected from household residents. At this point, logger-based studies begin to encounter the same issue of selective customer participation faced by all other data gathering efforts requiring customer consent and cooperation. For additional information about data loggers see Mayer, Peter.M., DeOreo, William.B., Opitz, Eva M. *et al*., *Residential Uses of Water*, A study funded by the AWWA Research Foundation, Denver, CO: AWWA Research Foundation and American Water Works Association, 1999.

[14] See, for example, *Water Conservation Baseline Study Final Report*, A study conducted by The Demand Management Company with Planning and Management Consultants, Ltd.; and *Market Saturation of Residential Water Conservation Devices*, A study conducted by Applied Management and Planning Group, August 1994.

[15] There are a variety of factors influencing the answers people give to survey questions. Research has shown that survey respondents often given *socially desirable responses* (i.e., answers that they consider to be socially acceptable) and that they will answer questions about which they have little or not knowledge. For a good but somewhat dated discussion of these types of issues, see Schuman, Howard and Presser, Stanley, *Questions and Answers in Attitude Surveys*, New York: Academic Press, 1981.

efficient showerheads and toilets from their less efficient counterparts based on visual inspection alone.

On-Site Inspections

Residential water "surveys", or "audits," have been a core element of many utility conservation programs for more than a decade. Often publicized using bill inserts, public service announcements and paid advertisements, water use surveys generally consist of a combination of customer education, leak detection, device flow measurement and plumbing device retrofits performed by utility paid personnel.[16] Although these types of surveys may seem a reasonable basis for estimating device saturation, households participating in residential surveys often differ in important ways from the "typical" household.[17] Using data gathered during the course of standard residential surveys to estimate the saturation of devices among all households—including non-participants—can therefore be problematic.

Method Used for the Orange County Saturation Study

The method used to collect device saturation information for the *Orange County Saturation Study* attempted to combine the advantages of the telephone and on-site inspection data collection methods while minimizing their respective disadvantages.

D. Research Design

1. Sample Design

The hypothesis that low-flow showerhead retrofits have occurred in a geographically uniform manner in Southern California was tested by comparing saturation rates estimated from a random sample drawn from all of Orange County with an independent sample drawn from the combined cities of Buena Park and Fullerton. The first step in developing an appropriate research design for generating the data needed to test this hypothesis involved calculating the size of the two samples needed to satisfy the statistical criteria prescribed by the urban *MOU*. Second, suitable sampling plan and data collection strategies needed to be developed. These matters are discussed below.

[16] While most such surveys are conducted by utility personnel or paid inspectors, some innovative programs have used school children to conduct these types of surveys.

[17] Many utilities, for example, target these types of surveys at their largest customers. Even without such selective targeting, research suggests that those who volunteer to participate in programs like these tend to be systematically different from members of the general population of interest. They may, for example, be more conservation-oriented, or their economic situation may make them more inclined to take advantage of "free" offers.

2. Sample Size Requirements and Sample Targets

Based on the sample size calculations described in Appendix A, approximately 75 dwelling units are required to estimate LFSH saturation with 95% statistical confidence and an error margin of ±10%. To allow for missing data and sample design effects, the target sample size was increased to 100.[18] Accordingly, testing the geographic uniformity hypothesis requires a random sample of 100 SF and 100 MF dwellings from the pre-1992 Orange County housing stock (***county sample***) and a similarly sized sample of SF and MF pre-1992 units drawn from the combined cities of Buena Park and Fullerton (***city sample***).[19]

The size of the county sample was governed more by budget restrictions than by statistical criteria. Together, Metropolitan and the Municipal Water District of Orange County (MWDOC) agreed to jointly fund the cost of adding 600 households to the county sample.[20] The augmented sample was designed to produce a representative sample of the current Orange County housing stock. Based on California Department of Finance housing statistics for the year 2000, it was expected that the augmented sample would consist of roughly 65% SF and 35% MF housing units. The augmented sample was also expected to include approximately 8% SF and 8% MF units constructed during or after 1992. Table 1 contains a breakdown, by dwelling type and by age, of the sampling targets for the city and the county samples.

Table 1 City and county sample targets, by dwelling type and age of structure

Dwelling Type	City Sample		County Sample	
	Built before 1992	Built during and after 1992	Built before 1992	Built during and after 1992
Single-family	100	0	359	31
Multi-family	100	0	196	14
Total	200	0	555	45

3. Sampling Plan and Data Collection Strategies

This section describes the sampling plan and data collection strategies used in generating the county and city samples.

Developing the Sampling Frame

A ***sampling frame*** can be thought of as a master listing of all members of the population of interest. The sampling frame plays two important roles in survey sampling. First, it serves as a mechanism for identifying and selecting sample participants. Second, it is an important point of

[18] See Appendix A for a discussion of sample design effects.
[19] The sample used to test the geographic uniformity hypothesis was restricted to pre-1992 dwellings because BMP 2 only requires showerhead retrofits in pre-1992 units. This is because all new housing constructed during 1992 and later is assumed to be fitted with low-flow showerheads in accordance with state and federal law.
[20] A grant from the United States Bureau of Reclamation also enabled the collection of additional survey data.

reference for assessing how well the final sample represents the population. Customer billing files are often used as the sampling frame for studies involving utility customers.

The type of information available in the typical utility's customer billing system usually depends on the way the service is metered. Conventional single-family dwellings in Southern California virtually always have dedicated water meters. In these cases, the billing system will generally include the customer's name, telephone number, and address. Apartment buildings, on the other hand, are usually master-metered. In these cases, the billing record usually contains the name and contact information for a non-resident owner or property manager. A second limitation is that the billing record seldom includes the number of separate residential units housed in the master-metered unit. Therefore, while the billing file may serve as a suitable frame for drawing a sample of single-family households, the fact that it does not separately enumerate units within master-metered buildings means that the utility's customer billing system is, by itself, not a suitable frame for drawing a representative sample of all customers served by the utility.[21]

Aside from the limitations already discussed, it would have been logistically impossible and prohibitively expensive to collect and combine customer billing records from all of the retail water agencies serving Orange County. An alternative method was therefore needed to enumerate households for the city and county samples.

Telephone Screening Survey

The sample of households for the *Orange County Saturation Study* was developed using a telephone screening survey. The first step in the process involved randomly generating a list of approximately 62,000 telephone numbers. [22,23] For the county sample, this master list of telephone numbers was randomly divided into 12 equally sized subsets, or *replicates*. One replicate was randomly selected and used to begin the telephone screening survey.[24] Each telephone number in the replicate was called a maximum of seven times. The time and day of calls made to each number were varied systematically to maximize the chances of contacting a household resident. Messages were not left on answering machines until the fifth attempt. From that point on, a brief message was left stating that the call was being made on behalf of a public agency and that the household had been randomly selected to participate in an important study. The message also said that a final contact attempt would be made during the next several days.

[21] Building on the address information contained in the billing system, a sophisticated multi-stage sampling design involving field enumeration and random selection of units within master-metered structures could produce a representative sample of the utility's entire customer base. This type of design is, however, expensive, time consuming and logistically difficult.

[22] Briefly described, such lists are generated by identifying all assigned telephone exchanges (i.e., the first three digits in a normal eight digit number), screening out exchanges that are reserved by the telephone company for non-residential use, then randomly generating the last four digits to form a complete local number.

[23] The list of telephone numbers was purchased from a company that specializes in developing such lists for use in survey sampling.

[24] The use of replicates helps ensure proper proportionate coverage of the area being sampled.

The purpose of these messages was to encourage people who use their answering machines to screen calls to pick-up the phone upon hearing the reason for the call.

If contact was not made by the seventh attempt, the telephone number was dropped from the sample and randomly replaced with another number from the same replicate. This process was repeated until all numbers in the replicate had been called up to seven times. After that, a new replicate was randomly selected and the entire process was repeated. This was done until 800 eligible households were identified and recruited as study participants.[25] The telephone interview was conducted in both Spanish and English.[26]

The telephone screening survey used to enumerate and recruit households for the city sample followed the same basic protocol with several minor adjustments. Contacted households were screened to ensure they were located in Buena Park or in Fullerton. Additionally, quotas were used to ensure that the resulting sample contained enough single- and multi-family households to satisfy the statistical requirements for testing the LFSH saturation hypothesis.[27]

Interviewing and Recruiting Study-Eligible Households

A carefully developed script was used to encourage participation in both the telephone and the on-site inspection phases of the study. After confirming that the randomly dialed telephone call had reached a residence in Orange County (county sample) or in Buena Park or Fullerton (city sample), prospective participants were told the study was sponsored by two public water agencies. They were also told that the telephone interview would only take a few minutes and that the information collected would be used in planning for Orange County's future water needs. Those who hesitated were given additional information about the importance of water

[25] Although the study design provided for completing 600 inspections for the county sample, a total of 800 study-eligible households were identified during the telephone enumeration survey. The extra 200 households were used as replacements for participants who subsequently reneged on their initial agreement to participate in the on-site inspection phase of the study.

[26] Consideration was initially given to conducting the screening interview in additional languages. According to 1990 Census figures, 10% of the population in Orange County reported not being fluent in English. After Spanish, Vietnamese, spoken by roughly 3% of the county population, was the third most common language. But the cost of interviewing in this and other languages spoken in Orange County was prohibitive and could not be justified given the few non-English and non-Spanish speaking households that would actually be captured by the sample. Furthermore, initial concerns about the potentially detrimental effects of systematically excluding certain households due to language barriers was assuaged after consulting with several individuals having extensive experience conducting telephone interviews in Orange County. According to these authorities, most non-native speaking households in the county include at least one English-speaking member. As it turned out, language barriers prevented telephone interviews from taking place for only 1.8% of all telephone contacts.

[27] The sampling targets were established to satisfy the statistical precision requirements specified by the urban *MOU*. The use of these targets, and the fact that the sample was not designed to yield proportionate representation of single- and multi-family housing units within the two cities, means that the city sample cannot be used to draw generalizations about the saturation of specific devices for all households comprising the two cities. In contrast, the sampling design used for the county sample does permit generalizations to be made to the full pre-1992 Orange County housing stock.

resource planning in Southern California and about the linkage between that effort and the study in which they were being asked to participate. Respondents who questioned the legitimacy of the study were given the name and telephone number of a person at the MWDOC to contact if they wanted to verify the authenticity of the study or if they had additional questions.

All telephone survey participants were asked a brief set of questions about tenure, dwelling type and age, length of residence, and the number of full-time residents. They were also asked if the residence received and paid a water bill and whether any showerheads or toilets had been replaced since 1992. These questions were asked for several reasons. First, it was anticipated that some households would be willing to participate in the telephone interview but not the on-site inspection phase of the study. Collecting the same basic information from all interviewed households would later serve as the basis for determining if the group of households who participated in the on-site inspection component of the study differed in important ways from those who declined to participate. Second, asking the same set of questions twice—once during the telephone interview and again during the on-site inspection visit—provided the information need to empirically assess the reliability of respondent self-reports.[28]

At the conclusion of the telephone screening survey, study-eligible households were asked to participate in the on-site inspection phase of the study. They were told that the inspection visit would take 30 minutes and that all households who completed the in-home portion of the study would be entered into a lottery for a $500 gift certificate.[29] Those agreeing to participate were asked for the best day and time for a technician to contact them to schedule the on-site visit. The telephone survey was conducted between September and November of 2000.

The effort required to identify and recruit the final sample of study households can be summarized as follows:

- A total of 42,239 distinct telephone numbers were dialed a total of 167,952 times.
- These calls produced 2,802 telephone interviews. Of these, 1045 said they would participate in the on-site inspection phase of the study.
- An average of 161 calls were required to successfully recruit each study participant.

Appendix B contains a detailed description of the procedures used in conducting the telephone screening survey and the final call disposition of all 167,952 telephone calls. Appendix C contains a copy of the telephone survey instrument.

[28] If deemed sufficiently valid and reliable, using information collected from telephone surveys alone might prove acceptable, especially when balanced against the considerable expense of conducting on-site inspections.

[29] The lottery served two purposes. The first was to overcome people's natural reluctance to allow a stranger into their home to inspect their bathrooms and other plumbing devices. The second was as an incentive to retain as many as possible of those who had agreed to participate until they were contacted to schedule the on-site inspection visit.

Scheduling On-Site Inspections

Approximately every three days the names, telephone numbers and addresses of newly recruited study households were sent to the field inspection team. The inspection team then attempted to contact each household by telephone to schedule an appointment for the inspection visit. This proved unexpectedly difficult. Additionally, the proportion of households who subsequently refused to participate when contacted by the inspection team was also higher than expected.[30]

Appointment confirmation letters, printed on utility letterhead, were mailed once the inspection visit was scheduled. These letters reiterated the importance of the study. They also provided the name and number of a person at MWDOC to contact if people wanted additional reassurances about the study's legitimacy. Finally, the letter reminded people that they would be eligible for participation in the lottery after completing the on-site inspection. The on-site inspections were conducted between October and December of 2000.

On-Site Inspection Surveys

The on-site inspections averaged 30 minutes. Upon first arriving for a scheduled appointment, inspectors verified and corrected, as necessary, the address and dwelling type information collected during the telephone screening survey. Some questions asked during the telephone survey were repeated and several new questions were asked about household water use. Next, the inspector asked to see the bathroom "most often used" by household occupants. The presence or absence of toilets, faucets and showers was noted. Showerhead flow rates were measured, the toilet was classified based on its rated flush volume, and the presence of leaking showers, faucets and toilets was recorded. Next, the host was then asked if, and when, the showerhead and toilet had been replaced and, in the case of replaced fixtures, whether the device was obtained through a utility-sponsored program. Finally, in an effort to identify conservation program *freeriders*,[31] residents who said they had obtained their current showerhead(s) or toilet(s) through utility-sponsored programs were asked if they would have replaced the device, at their own expense, if the program had not existed. The entire process was repeated for the "next most frequently used bathroom" until all bathrooms in the dwelling were inventoried.

The effort required to complete the on-site inspection component of the study can be summarized as follows:

[30] There is some evidence that the telephone survey respondent's initial agreement to participate in the on-site component of the study was sometimes subsequently overruled by someone else in the household. This suggests that it may be useful for future studies to attempt to identify and obtain authorization for the on-site visit from a "key" decision-maker within the household.

[31] A program *freerider* is one who would normally have replaced a showerhead or toilet at his/her own expense (for example, due to device failure or remodeling), but who opportunistically obtained a utility-subsidized device. The effect of program freeriders is to reduce program cost-effectiveness by using program funds to help subsidize a purchase that would otherwise have been made at the homeowner's own expense.

- Among the 1,045 households for which on-site inspection visits were successfully scheduled, 762 inspections were actually completed.

- Among those who reneged on their initial agreement to participate, most did so during the telephone contact made to schedule the visit. A handful scheduled the visit but declined to participate when the plumbing inspector arrived at the appointed time.

- For the county sample, on-site inspections were completed for 385 single-family and 196 multi-family dwellings. For the city sample, 87 single-family and 94 multi-family dwellings were inspected.[32]

Appendix C provides additional information about the on-site inspections, including a detailed breakdown of the final disposition of all inspection scheduling attempts. Appendix E contains a copy of the data collection instrument.

4. Strategies for Maximizing the Reliability and Validity of Key Study Variables

Because of their critical importance, special attention was given to the procedures used to measure three key study variables: household dwelling type, construction date, and the presence of water conserving showerheads and toilets. The issues involved and the procedures used are discussed in this section.[33]

Dwelling Type

One of the main study goals was to develop separate LFSH saturation estimates for single- and multi-family units. The advantages of using the Department of Finance's definition of single- and multi-family dwelling units have already been discussed. Effectively realizing these advantages, however, required that classification of study households be both reliable and valid.

[32] Although the number of completed inspections was slightly less than originally planned, this did not adversely affect the results of the statistical tests discussed in Chapter IV.

[33] As discussed previously, the correct classification of dwelling units was important for at least three aspects of the survey. First, there was an interest in comparing device saturation estimates for these two types of dwellings. Second, the proportion of single- and multi-family households enumerated during the telephone screening survey was compared, on an ongoing basis during the course of the survey, to benchmark California Department of Finance (DOF) housing statistics. The purpose was to identify any emerging discrepancies between the results of the enumeration survey and these benchmark statistics, and to take necessary corrective action, should the need arise. Finally, the study design called for the use of DOF housing statistics in developing sampling weights to be used in adjusting for any post-survey imbalances between the single- versus multi-family composition of the on-site inspection sample and the "true" proportion as these two dwelling unit types. For this purpose, the DOF housing statistics for Orange County were assumed to represent the true population proportions

As discussed previously, telephone survey respondents were asked several questions intended to provide the information needed to classify dwellings using Department of Finance criteria.[34] Although the questions used for this purpose during the telephone interview seemed simple and unambiguous,[35] it was anticipated that some respondents would provide erroneous information, causing their housing unit to be incorrectly classified. Site inspectors were therefore instructed to verify the dwelling classification during the site visit. Since inspectors were trained to identify and properly classify dwelling units using DOF criteria, the resulting classifications satisfy the study's reliability and validity requirements.[36]

Construction Date

Since 1992, state law has restricted the sale and installation of showerheads, toilets and aerators to water-efficient models. Date of construction therefore played an important role in the saturation study. This is because non-conserving showerheads should only be found in pre-1992 units. Accordingly, CUWCC's residential plumbing retrofit requirement only applies to pre-1992 units.

All telephone survey respondents were asked how long they had lived at their current residence. Those saying they had resided in the unit for 10 or fewer years were asked if the building was built before 1992. It was assumed that most owners would know when their home was built, but it was also recognized that the memory of some owners could prove faulty. Furthermore, renters may have little reason, or opportunity, to know when their dwelling was built. For these reasons, it was anticipated that the answers given to the question about construction date asked during the telephone survey could be even less reliable than these given in response to the question asked about dwelling type. Accordingly, the study design provided for the use of a proprietary database, containing information compiled from the Orange County Assessor's Office, to determine the actual construction date of all inspected units.[37]

The parcel database turned out to be more difficult to use than originally anticipated. The first problem was the difficulty of matching street address information for the study households to the street address information in the parcel database. While this will come as little surprise to anyone who has tried matching separate data files using street addresses, the number of non-matches was quite high. Only about 60% of all study household addresses could be successfully

[34] California Department of Finance, *California Statistical Abstract*, Table I-3, footnote a, October 2001. This table can be accessed via the Department's web site at www.dof.ca.gov/HTML/FS_DATA/stat-abs/tables/i3.xls.
[35] See survey items J and K in Appendix C.
[36] Several inconsistencies were noted between the dwelling unit classifications developed from the telephone screening survey and those returned during the early on-site inspection surveys. A supervisor, sent out to resolve these inconsistencies, found that in virtually all cases, the classification assigned by the site inspector was correct. A few incorrect classifications were noted, however, leading to some remedial training and closer supervision of the on-site inspectors. While the possibility therefore exists that some small percentage of units are misclassified, the inspector classifications are nonetheless believed to be highly reliable.
[37] Information about the Metroscan® parcel database can be found at ww.firtanres.com/thml/contact.html.

matched to the parcel database.[38] The Orange County Assessor's Office was able to provide construction date information for about 31% of the remaining households. Construction dates for the remaining 9% of inspected households were estimated by combining information about length of tenure with answers to questions about the date of construction asked during both the telephone interview and the on-site inspection visit. The composite construction date variable that was formed from these separate information sources is considered highly reliable and is used to classify pre-1992 units in all subsequent analyses.

Classifying Showerheads and Toilets

As discussed earlier, low-flow showerheads cannot be reliably identified based on visual inspection alone. For this reason, plumbing inspectors conducted flow rate measurements on all showerheads. A two-step procedure was used to measure the volume of water flushed by all inspected toilets. Inspectors first flushed the toilet and watched for the distinctive flushing behavior of an ultra-low flush (1.6 gallon) device. If the results were indeterminate, the volume of water flushed by the device was measured directly.

Appendix F contains a more detailed description of the measurement procedures and protocols used during the on-site inspections.

E. Coordinating Data Collection Activities

Monitoring and coordinating of the various data collection activities was logistically challenging. Information about willing, study-eligible participants identified during the screening survey had to be relayed to the field inspection team. They then had to contact and schedule the inspection visits. Once scheduled, appointment confirmation letters were sent. Contacts at all retail utilities in Orange County and the MWDOC field coordinator received complete lists several times a week containing the names and telephone numbers of all households contacted by the telephone survey company. This information was needed in the event that customers contacted the utility

[38] The on-site inspectors found discrepancies between the actual street address and the address recorded during the telephone screening survey for 13.7% of all households inspected. Since many of these involved misspellings, the telephone interview protocol was modified to ask the respondent to spell the street name rather than leaving the telephone interviewer to guess the correct spelling. It is strongly suggested that this procedure be used in any future telephone surveys. It is also suggested that interviewers explicitly ask if there is an apartment or unit number and if the street name includes a suffix (e.g., Ave, Blvd., etc.). Finally, although it will add significantly to the time required for conducting the interview, it may be worthwhile to have the interviewer repeat the full street address and ask the respondent to confirm that it has been correctly recorded.

to verify the authenticity of the study or to express other concerns. The study project director needed current information about telephone contact rates, participant recruitment, and the number and benchmark characteristics of completed on-site inspections for both the county and the city samples. Necessary adjustments needed to be expeditiously communicated by the project director to the telephone survey company and to the field inspection team.

Appendix G presents a detailed discussion of study implementation issues and the methods used to address these issues and to coordinate the overall data collection and monitoring effort.

III. DATA QUALITY

Evaluating the quality of data obtained from a survey sample involves at least two basic questions. First, how well does the sample represent the population from which it is drawn? Second, how reliable and how valid is the information that was collected? To the extent minor deficiencies are discovered, remedial measures can be employed. In the case where such corrections are not possible, appropriate caveats need to be developed and used to qualify study results.

This chapter examines the geographic representativeness of the households comprising the *Orange County Saturation Study* survey sample. The reliability and validity of measures obtained from the study households is also assessed. With one notable exception, discussed below, the overall quality of the sample seems quite good. Some weighting is necessary to correct for imbalances between the spatial distribution of study households and the distribution of the Orange County housing stock. But the size of this imbalance is so small that differences between weighted and unweighted sample statistics are trivial in most cases.[39]

As discussed in Chapter II, a subset of survey questions was asked twice—once during the telephone survey and again during the on-site inspection. The purpose was to generate the information needed to empirically estimate the reliability of respondent self-reports. The results, discussed in Section III.D, indicate that respondent self-reports are not a suitable basis for reliably estimating the residential saturation of devices like low-flow showerheads and ultra-low flush toilets. These findings also raise some potentially troublesome questions about the reliability of the answers people give to simple, seemingly innocuous questions about basic household characteristics and other factual matters.

A. The Exclusion of Post-1991 Multi-Family Dwellings

The number and type of households recruited by the telephone screening survey was closely monitored throughout the survey. So was the number and type of completed household inspections. This allowed the early identification of any emerging discrepancies between the expected share of single- and multi-family dwellings in the two samples and the implementation of any necessary corrections. Based on the information available at the time, these few sampling adjustments that were implemented appeared quite successful at improving the overall representativeness of the county sample.

Since the city sample was designed to include only pre-1992 dwellings, only those units reported as being built before 1992 during the telephone screening survey were eligible for inclusion in

[39] The same is true for sample design effects. While the showerhead saturation hypothesis tests in Chapter IV explicitly account for design effects, the effects are so small that the conclusions drawn from these tests are the same regardless of whether or not the statistics are adjusted for design effects. A similar result is for the descriptive analysis discussed in Chapter V. In that case, inclusion or exclusion of sample weights has no apparent effect on the main substantive conclusions developed in that chapter.

the city sample. While it was expected that faulty memories and flawed knowledge would result in some wrong answers to the construction date screening question, it was anticipated that these types of response errors would tend to be random. In other words, it was expected that errors from those incorrectly reporting a post-1991 construction date would offset by those incorrectly reporting that their dwelling was built before 1992. Since random response errors of this type would not systematically bias the final sample, and since it was thought that the true construction date of all inspected households would eventually be known, misreported construction dates were not viewed as a threat to the overall representativeness of the final city sample. The main concern was therefore ensuring that the final sample of inspected units of verified pre-1992 vintage was sufficient to meet the sample size requirements for the city sample.[40]

Because the county sample was designed to produce a representative sample of the entire year 2000 Orange County housing stock, construction date was not one of the variables used in monitoring the enumeration of this sample. Rather, it was expected that a spatially representative county sample would yield a proper mix of pre-1992 and post-1991 study households. When reviewing the construction date of all inspected households, however, it was discovered that the county sample contained no post-1992 multi-family units.[41] Despite subsequent investigations, the reason for this outcome remains a mystery. The fact that the telephone screening survey identified and successfully recruited the expected share of post-1991 multi-family units suggests that a problem may have occurred during the inspection phase of the study. The procedures used for scheduling and conducing the on-site inspections, however, make it difficult to understand how this systematic exclusion of post-1991 multi-family units could have happened.

Aside from the exclusion of post-1991 multi-family units, it is important to note that the quality of the county sample seems quite good in all other respects. Accordingly, the primary effect of the omission is that sample-based inferences and generalizations involving multi-family units will need to be limited to the current stock of pre-1992 households. In practice, this restriction is of little consequence since the plumbing retrofit activity of interest in this study involves the pre-1992 housing stock.[42]

[40] Sample size requirements are discussed in Section II.D.2 and Appendix A.

[41] This result was discovered only after attempting to determine the true construction date of all study households. This occurred approximately four months after the end of the data collection process. Budget constraints and expired agreements made it impossible to undertake a supplementary sample of multi-family dwellings.

[42] Post-1991 units would not have been included in the saturation analysis of LFSHs and ULFTs since BMP retrofit requirements only apply to pre-1992 dwellings. This is because by code, all housing units constructed after 1991 were required to be fitted with low-flow showerheads. Since the building code requirement governing installation of ULFTs was phased in between 1992 and 1994, the availability of post-1991 multi-family units may have allowed a slightly refined estimated of ULFT saturation. Analysis conducted using single-family units, however, showed that changing the implementation year from 1992 to 1994 produced no detectable changes in the estimated saturation rate.

B. Geographic Coverage

From a sampling perspective, the county sample had two main goals. The first was to produce a geographically representative sample of Orange County housing units. The second was to produce a sample with proportionate representation of single- and multi-family dwellings. Because of the omission of post-1991 multi-family units discussed above, however, the criteria to be used in assessing the representatives of the county sample need to be modified slightly. The first now involves the question of how well the geographic distribution of inspected single-family units compares with the geographic distribution of all year 2000 single-family units in Orange County. The second involves the question of how well the composition and distribution of the subset of inspected pre-1992 units matches the existing stock of pre-1992 housing units.

Appendix H contains information that can be used to answer both questions. As this appendix shows, the proportion of study housing units within each sample stratum (i.e., year 2000 single-family, pre-1992 single-family and pre-1992 multi-family) compares quite favorably with the "expected" proportion of each housing segment.[43]

C. Sample Weighting

Sample weights were developed to correct for observed imbalances between the sample results and the population benchmarks. The resulting weights are small and well within conventional bounds.[44] Indeed, the statistical results discussed in Sections IV and V appear largely insensitive to the use or to the exclusion of sample weights. However, in the interest of technical accuracy, the sample weights will be used in all subsequent analyses.

D. Selection Bias

Any study that depends on the voluntary participation must consider the possibility of *self-selection* and its associated statistical consequences.[45] It would not be surprising, for example to

[43]Chi-square goodness of fit tests indicate that the discrepancies between observed and expected proportions are not statistically significant at the 0.05 level for multi-family units and at the 0.01 level for single-family units. The p-values for these tests are actually smaller than what would be expected based on the seemingly close correspondence in most cases between observed and expected proportions of households assessed across cities. In this regard, the chi-square test results may be untrustworthy for several reasons. First, expected cell frequencies fall below 5 in several cases and observed frequencies equal zero in several others. Both situations are inconsistent with important assumptions where applications of the chi-square statistics to data like these are involved. Second, the test has a relatively large number of degrees of freedom. In such cases, the resulting power of the test can cause relatively minor differences between observed and expected frequencies to yield test results indicating that the differences are statistically significant. Placed in this context, the p-values obtained in the goodness of fit tests seem quite respectable and suggest that observed differences may be the result of normal sampling error.

[44] See Appendix A.

[45] Self-selection is one of the major reasons caution needs to be exercised in using information gathered from participants in utility-sponsored residential surveys as a basis for drawing inferences about the saturation of water-efficient devices among the larger population of utility customers. In particular, there are good reasons for suspecting that these types of programs tend to draw disproportionately from those who are more interested than the

find that people with an interest in conservation are more willing to participate in a conservation-related study than are "typical" members of the general population. The effect of this type of selective participation will be that the survey results will disproportionately reflect the opinions of those for whom conservation is a salient issue. To the extent that these opinions differ from those held by the more general population of interest, sample-based inferences will produce biased estimates of the population parameters of interest. [46]

Considerable effort was made to maximize participation of all contacted households in the telephone screening survey. Up to seven attempts were made to contact a member of each randomly selected household with the day and time of successive calls being systematically varied in an effort to maximize the chances of finding a resident at home. Messages were left on answering machines beginning with the fifth attempt stating that a local water utility was calling and that another attempt would be made during the next several days. The objective was to encourage individuals who use their answering machine to screen calls to pick-up the phone. A carefully worded script was used to let contacted households know the telephone interview would only take a few minutes and to solicit their participation by appealing to both their civic and their self-interest. A review of the geographic distribution of households participating in the telephone survey suggests that these strategies successfully produced a geographically representative sample of Orange County households.[47]

The possibility of systematic differences between households who agreed and those who declined to participate in the on-site inspection phase of the study is a second potential source of selection bias. A series of probit regression models were developed to examine this possibility. The results do not indicate any important differences between participating and non-participating households.[48]

---

typical customer in saving water or in obtaining free or low cost plumbing devices. In an attempt to minimize this type of self-selection, the telephone screening survey script did not explicitly state that the study involved water use efficiency. Rather, prospective participants were told that the information gathered would be used by local utilities to meet the water needs of Southern California over the next 20 years. It is believed that this veiled appeal to people's self- and civic interests would result in the successful recruitment of a broader segment of the population than one that emphasized conservation.

[46] Technically speaking, selection bias can adversely affect both the internal validity and the generalizability of inferences based on information gathered from a selective segment of the full population. See Berk (1983) for a readable treatment of these topics. See Muthen and Yang (1990) and Hollis (1991) for more technical developments.

[47] Geographic representativeness is, of course, not a sufficient basis for ruling out the possibility of selection bias. A more rigorous assessment would include a comparison of the social and economic characteristics of participants and non-participants. While early consideration was given to collecting this type of information during the telephone survey, it was decided that asking questions about household income, education, race/ethnicity and similar variables would increase the non-participation rate by raising suspicions about the "real" motives of the study. This is because most people would fail to see the relevance between this type of personal information and a study that had represented itself as being focused on the types of water-using devices people have in their homes. Furthermore, it was anticipated that even if asking these types of questions did not result in outright refusals to participate in the on-site inspection phase of the study, the inevitable combination of unreliable answers and missing data would have rendered the information collected largely unsuitable for its intended purpose.

[48] See Appendix H.

In conclusion, the available evidence suggests no reason for believing that participant self-selection—i.e., systematic differences between those who participated and those who refused to participate in either the telephone enumeration survey on the on-site inspections—is a major source of concern in this study. Combined with the good geographic coverage of the sample, the sample would seem to support generalizing study findings to the larger population of all pre-1992 housing units in Orange County.[49]

E. The Reliability and Validity of Self-Reported Information

The research design developed for the *Orange County Saturation Study* was motivated largely by the belief, discussed earlier, that the saturation rate of low-flow showerheads and ultra-low flush toilets could not be reliably estimated using information collected during a telephone survey. The material discussed in this section clearly demonstrates that concerns about the reliability of household self-reports were well founded.

1. Dwelling Unit Classification

As previously discussed, the criteria used by the California Department of Finance to classify dwellings are relatively straightforward. Single-family dwellings consist of freestanding structures and attached dwelling units separated by a common wall extending from the ground to the roof. Multi-family dwellings have a separate residence above and/or below each separate residential unit.[50]

Two questions were used during the telephone screening survey to classify dwellings. The first distinguished between freestanding and multi-unit structures. Those who said they lived in a multi-unit structure were then asked if there was a separate residence located directly above or below the one in which they lived.[51]

It was possible to "trap" inconsistent answers given during the telephone interview in one particular instance—when the respondent said he or she lived in a freestanding detached single-family building but subsequently said there was a separate residence above or below the one in which they lived.[52] Twenty out of the total of 2,830 telephone survey respondents gave these types of inconsistent answers. While this number is relatively small, it does raise several questions. First, some degree of error in self-reported information appears inevitable, even when the questions asked seem quite simple and unambiguous. Perhaps respondents inadvertently

[49] Strictly speaking the study findings are limited to the pre-1992 housing stock as it existed during the fall and winter of 2000 when the sample was drawn.
[50] Multi-family dwellings consist largely of apartment buildings. But they also include high-rise condominium developments.
[51] See questions J and K in Appendix C for the exact wording of the questions used.
[52] In the case of such inconsistencies, the telephone interviewer asked Q.K a second time in an attempt to resolve the inconsistency.

reversed the response options or found the question confusing. Perhaps they were not sufficiently engaged by the interview and simply reacted to the first response option presented to them. Perhaps they were, for whatever reason, intentionally deceptive. Regardless of the exact reason, this result shows that answers to what may seem to be simple questions about objective, non-personal and non-threatening questions are likely to suffer from some degree of response error.

Response error was affected the initial classification of dwelling units. Among the 762 housing units visited by trained inspectors 81, or 10.6%, were found to have been misclassified based on respondent-reported information collected during the telephone survey.[53]

2. Date of Construction

Since 1992 the sale and installation of showerheads, toilets and faucet aerators has been restricted by law in California to water-efficient models, date of construction plays an important role in device saturation studies. Since utility-funded retrofit programs target pre-1992 units, estimates of the remaining stock of water-inefficient devices, and of the rate at which these units are naturally replaced, are of critical interest to conservation program administrators.

All telephone survey respondents were asked how long they had lived at their current residence. Those saying they had resided in the unit for 10 or fewer years were asked if the building was built before 1992. Information obtained from the Orange County Assessor's Office was later used in establishing the true construction date of inspected units.[54]

One measure of the reliability of respondent-reported construction dates can be developed from respondents who said they had lived at their current residence for 10 or fewer years. This subset of respondents was then asked if the unit in which they lived had been built before or after

[53] Inconsistencies were first noted between the dwelling unit classifications developed from the telephone screening survey and those returned during some of the early on-site inspection surveys. A supervisor, sent out to resolve these inconsistencies, found that it several cases the classification assigned by the on-site inspector was wrong. This led to some remedial training and closer supervision of the on-site inspectors. The possibility therefore remains that some small share of the total number of misclassifications resulted from inspector error. Nonetheless, the inspector's classification is taken as definitive for purposes of this study.

[54] The assessor's records did not contain parcel information for approximately 9% of all inspected households. In these cases, construction date had to be estimated by combining occupant-reported information about length of tenure and about construction date. Crosschecks indicate the resulting estimates are highly reliable in most cases.

The inability to match these inspected units to assessor parcel records may have resulted, in part, from the fact that on-site inspectors found discrepancies between the actual street address and the address recorded during the telephone screening survey for 13.7% of all inspected. Since many of these involved misspellings, the telephone interview protocol was modified to ask the respondent to spell the street name rather than to leave the telephone interviewer to guess the correct spelling. It is strongly suggested that this procedure be used in any future telephone surveys. It is also suggested that interviewers explicitly ask if there is an apartment or unit number and if the street name includes a suffix (e.g., Ave, Blvd., etc.). Finally, although it will add significantly to the time of the interview, it may be worthwhile to have the interviewer repeat the full street address and ask the respondent to confirm that it has been correctly recorded.

1992.[55] Of the 543 people answering this question, 13.6% answered incorrectly or said they did not know the answer.

3. LFSH and ULFT Device Saturation Estimates

Table 2 compares estimated saturation rates for low-flow showerheads and low-flush toilets based on respondent self-reports with actual measurements obtained during the on-site inspections. As this table shows, basing saturation rates on information provided by household occupants would have underestimated the true saturation rate of LFSHs by about 34% and the true saturation rate of ULFTs by 42%.[56]

Table 2: Saturation rates based on respondent self-reports versus on-site measurements

Saturation Estimate	Telephone Survey	On-Site Inspection	No. of Households Used in the Calculation
Low-flow showerheads	43.5%	64.7%	740
Ultra-low flush toilets	28.7%	48.7%	741

Note: Estimates are based on inspected households from both the city and county samples

4. Other Response Inconsistencies

Information about tenure, whether the household received and paid a water bill, number of full-time household residents, number of bathrooms, and the number of bathrooms with showers was collected both during the telephone interview and during the on-site inspection. Table 3 shows the percentage of inconsistent answers given on the two occasions.

Table 3: Reliability of self-reports about other basic household characteristics

Question	Inconsistent Responses	Don't Know Responses
Do you own or rent?	6.0%	3.2%
Does your household receive and pay a water bill?	12.1%	3.6%
How many people live in your home on a permanent basis?	15.2%	0.0%
How many bathrooms do you have in your home?	13.9%	0.0%

[55] This question was originally asked of this subset of respondents because it would not be possible to estimate pre-92 versus post-91 unit construction using years of residence alone. In retrospect, the question should have been asked of all respondents.

[56] A 1994 baseline study in Marin County, California, found that respondents tended to *over-report* the presence of LFSHs and ULFTs (Demand Management Company *et al.*,1994). Since the study was conducted during a drought, the over-reporting may have resulted from a heightened awareness of water conserving devices. Regardless of the actual reason, the contrasting results show that estimates based on customer self-reports can either over-estimate or under-estimate true saturation rates. The problem, of course, is that it is impossible to know the direction of bias based on the self-reported information alone.

Some of this inconsistency may have resulted from the fact that the answers were given by different people–i.e., the person hosting the on-site visit may have been different from the person who participated in the telephone-screening interview. Accordingly, the percentage of inconsistent responses probably overstates what would have resulted had the same person answered questions during both the telephone screening and the on-site surveys.[57] Still, the extent of discordant answers given to relatively simple, factual questions is surprising and suggests several cautions. The first is that the accuracy of information given by customers during surveys should always be viewed with a healthy dose of skepticism. The second is that when there is no alternative to relying on customer self-reports, a concerted attempt should be made to identify and talk to the most "knowledgeable" person in the household. Finally, the more esoteric the issue being studied (like whether or not bathrooms have been retrofitted with low-flow showerheads and ultra-low flush toilets), the less reliable the information obtained from household occupants will be.

[57] Additional factors may account for some of this inconsistency. These include slight differences in question wording and the different methods of asking the question (i.e., telephone versus in-person). Allowances also need to be made for the possibility that the questions were ambiguous to some respondents.

IV. ESTIMATING LOW-FLOW SHOWERHEAD SATURATION

A. Introduction

This chapter tests two hypotheses about the saturation of low-flow showerheads (LFSHs) in Orange County, California. The first is that this process has occurred uniformly throughout the county.[58] The second is that the 75% saturation level, signifying full implementation of BMP 2, has been achieved for both single- and multi-family households in the county.[59]

B. Testing the Hypothesis of Geographically Uniform Low-Flow Showerhead Saturation Rates

Table 4 summarizes the information needed to statistically test the hypothesis of geographically uniform LFSH saturation. Of particular interest is the saturation of LFSHs among single-family households in the city and the county samples and the corresponding contrast for multi-family units. Under the hypothesis of geographically uniform saturation, the estimated saturation rates should be statistically equivalent for comparable dwelling types.

Table 4: Geographic uniformity hypothesis test results, by sample and by dwelling type

Variable	County Sample		City Sample	
	SF	MF	SF	MF
No. of pre-1992 dwelling units inspected	358	196	85	93
No. of showerheads inspected	703	258	145	119
Mean LFSH saturation rate	66.9%	59.8%	66.2%	54.6%
Standard error of the mean	2.0%	3.3%	4.2%	4.8%
95% confidence interval width	±3.9%	±6.5%	±8.2%	±9.4%
95% confidence interval range	63.0–70.8%	53.3–66.3%	58.0–74.4%	45.2–64.0%

NOTE: Standard errors are corrected for intra-cluster correlation.

As Table 4 shows, the saturation rates estimated for the county and the city samples differ by less than 1% for single family households (66.9% v. 66.7%) and by about 5% for multi-family units (59.8% v. 54.6%). The question of whether these numerical differences are statistically significant is readily answered using their corresponding confidence intervals. Finding, for example, that the two saturation rates estimated for single-family units both fall within the region where the confidence intervals overlap would lead to the conclusion that the saturation rates estimated for the two samples are not statistically different. Conversely, finding that one of the estimated saturation rates falls outside the overlapping region would lead to the conclusion that the saturation of LFSHs is statistically different for single-family households in the city and

[58] The reasons for this conjecture are discussed in Section I.A. The basic testing logic and its requirements in terms of research design and data requirements are presented in Sections II.A and II.D.

[59] This hypothesis is tested following the criteria prescribed in the urban *MOU*. See Section I.A for a discussion of these requirements.

county samples. The first outcome is consistent with the hypothesis of geographically uniform LFSH saturation. The second outcome would cause the hypothesis to be rejected.

Figure 1 provides a visual representation of the information in Table 4. The double-headed arrows in this figure represent the 95% confidence intervals for the city and county samples. The vertical dotted lines define the interval where the two confidence intervals overlap. The "x"s in the figure correspond to the LFSH saturation estimates for single-family units in the two samples. Since both estimates fall within the region bounded by the vertical dotted lines, *the hypothesis of geographically uniform LFSH saturation among single-family housing units in Orange County cannot be rejected.*

Figure 1: Uniformity hypothesis test for single-family dwellings

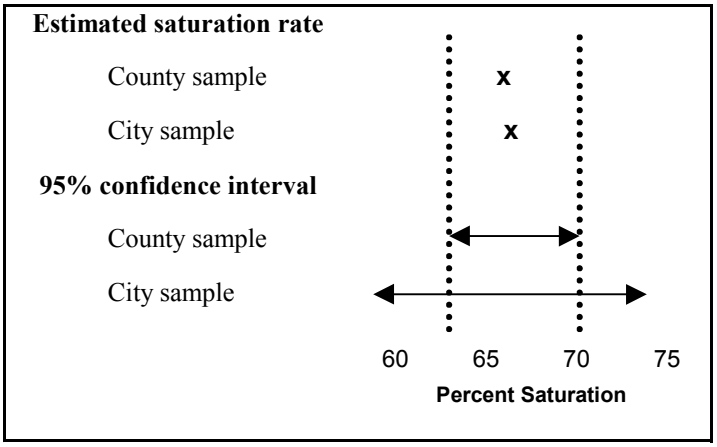
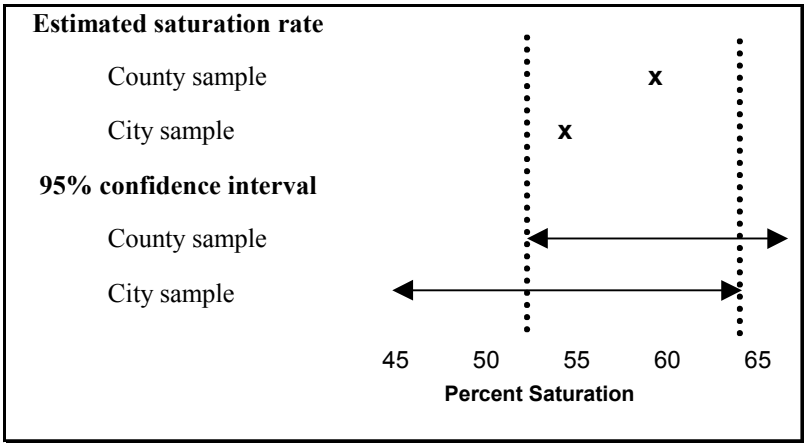


Figure 2 on the following page shows that the same result holds for multi-family units. Accordingly, *the hypothesis of geographically uniform LFSH saturation among multi-family housing units in Orange County cannot be rejected.*

Combined, these test results support the proposition that a county-level survey in Southern California can provide good estimates of the saturation of LFSHs among households served by retail agencies operating within the county.

Figure 2: Uniformity hypothesis test for multi-family dwellings



C. Testing the 75% Saturation Hypothesis

Based on the sample size calculations described in Appendix A, a random sample of approximately 100 showerheads is needed to satisfy the statistical criteria governing hypothesis tests involving the 75% saturation requirement for BMP 2. Accordingly, random sub-samples of approximately 100 showerheads from single-family units and 100 showerheads from multi-family units were drawn from the county and city samples. These sub-samples were used to calculate the summary test statistics in Table 5.

Table 5: Showerhead saturation test results, by sample and by dwelling type

Variable	County Sample	
	Single-family	Multi-family
No. of randomly selected pre-1992 dwellings	55	80
No. of showerheads	110	106
Mean LFSH saturation rate	66.6%	60.5%
Standard error of the mean	5.1%	5.1%
95% confidence interval width	±10.0%	±10.0%
95% confidence interval range	56.6–76.6%	50.5–70.5%

NOTE: Standard errors are corrected for intra-cluster correlation.

For single-family dwellings in both the county and city sub-samples, the 95% confidence interval captures, albeit narrowly, the required 75% saturation rate. Accordingly, a strict interpretation of the test results supports the proposition that the 75% BMP 2 coverage requirement has been met for single-family households in Orange County. For multi-family units, the 95% confidence interval fails to include the75% saturation threshold. As such, the hypothesis that the BMP 2 coverage requirement has been met for multi-family dwellings in Orange County must be rejected.

The test results described above were obtained by defining a low-flow showerhead as one having a maximum measured flow rate of 2.5 gpm. While this is the definition used in the 1992 plumbing code, Table 6 shows that relaxing the definition of a low-flow device to allow a maximum flow rate of 2.8 would lead to a different set of conclusions. Under this more liberal definition, the estimated saturation of LFSHs among single-family households would increase to 74%. The corresponding estimate for multi-family units would increase to 66%. From a narrow statistical perspective, this would mean that multi-family units in Orange County have also satisfied the 75% saturation criterion.

Table 6 : Showerhead flow rate distribution, by flow category and by dwelling type

Flow Category (gpm)	Cumulative Distribution, by Dwelling Type (county sample)	
	Single-Family	Multi-Family
<= 2.50	67.2%	59.9%
2.51-2.60	68.0%	59.9%
2.61-2.70	69.2%	60.7%
2.71-2.80	74.0%	66.0%
2.81-2.90	74.0%	66.0%
2.91-3.00	81.5%	73.3%
3.01+	100.0%	100.0%

The fact that conclusions about whether or not the MOU-prescribed 75% saturation level has been achieved are sensitive to relatively minor changes in the flow rate used to classify low-flow devices matters for several reasons. First, some degree of random measurement error in measuring flow rates is inevitable. As such, some share of showerheads with measured flow rates in the 2.71-2.80 gpm range could actually be mis-measured 2.5 gpm devices. Second, and more troubling, is the possibility that some devices labeled and sold as low-flow units may actually be manufactured with flow rates exceeding 2.5 gpm.

Without compelling evidence of pervasive measurement error, or that a substantial share of showerheads labeled as low flow devices are routinely manufactured with flow rates exceeding 2.5 gpm, there is no empirical basis for altering the initial test saturation test results and associated conclusions. Nonetheless, the number of showerheads with measured flow rates in the 2.51-2.80 range makes it difficult to dismiss the possibility that the study results may underestimate the actual saturation of manufacture-rated 2.5 gpm showerheads.

D. Conclusions and Recommendations

The results developed in this chapter paint a mixed picture. The hypothesis of geographically uniform saturation of low-flow showerheads among pre-1992 housing units has been strongly supported. Similarly, the available evidence consistently indicates that the required 75% saturation level has not been achieved for multi-family units as of the winter of 2000. Interpreting the results with respect to LFSH saturation among single-family units in Orange County is less straightforward.

Technically, the test results for single-family units are consistent with the hypothesis that this segment of the housing stock has attained the required 75% saturation level. But this interpretation is at odds with what would be concluded based on the estimates shown in Table 4. As that table shows, the 95% confidence interval estimated using the full county sample of single-family dwellings falls short of the 75% requirement. The difference results from the fact that the larger sample size used in estimating the statistics in Table 4 results in a much smaller confidence interval than the ±10% margin of error allowed under the *MOU* test criteria.[60]

[60] It is easy to show that the width of a confidence interval calculated for a given standard error and significance level is inversely related to the size of the sample.

V. RESIDENTIAL PLUMBING DEVICE AND USE PROFILE

A. Introduction

This chapter profiles the type and distribution of indoor water-using devices found in Orange County households. The saturation of ultra-low flush toilets, the incidence of leaking toilets and showerheads, and the incidence and frequency of use of dishwashers and clothes washers will be described. This information should be of interest to conservation planners and to those who use end-use models to forecast residential water demand. A limited attempt is also made to estimate the relative magnitude of *freeriders* among households who report having received a LFSH through a utility-sponsored program.

All tables in this chapter include statistics for the following segments of the Orange County housing stock: pre-2001 single-family ("Thru 2000"); pre-1992 single-family; pre-1992 multi-family; and the entire stock of pre-1992 housing units ("Full Sample, Pre-1992").[61] Because the pre-1992 housing stock is of special interest to conservation planners, most of the discussion focuses on this group.

B. Basic Dwelling and Household Characteristics

Table 7 summarizes basic characteristics of households comprising the county sample of inspected dwellings. As this table shows, single-family dwellings tended to have more bathrooms and more occupants. They were also more likely to be owner-occupied and to receive and pay their own water bills. The average tenure of those in single-family units was about 10 years, nearly twice as long as the typical occupant of multi-family housing.

Table 7: Basic dwelling and household characteristics, county sample

Variable	Single-Family		Multi-Family	Full Sample
	Thru- 2000	Pre-1992	Pre-1992	Pre-1992
Average number of bathrooms	2.4	2.3	1.4	2.0
Average showers per household	2.0	2.0	1.3	1.8
Percent receiving and paying a water bill	86.7%	87.0%	22.2%	64.1%
Percent owner-occupied unit	84.7%	84.6%	15.0%	60.6%
Average persons per household	3.3	3.3	2.8	3.1
Average length of residence, in years	9.9	10.4	5.0	8.5

[61] As discussed in Chapter III, the county sample does not include any post-1991 multi-family units.

C. Indoor Plumbing Devices

1. Showerheads

Flow Rates and Leaks

Table 8 compares the distribution of showerheads, by flow rate, for single- and multi-family units. As noted previously, the saturation of low-flow showerheads among pre-1992 units is higher for single-family units. Conversely, the incidence of leaking showerheads is greater among multi-family units. It is expected that the difference would be even more pronounced if separate estimates were calculated for apartment units. In other words, the saturation of

Table 8: Showerhead flow rates and leaks

Variable	Single-Family		Multi-Family	Full Sample
	Thru 2000	Pre-1992	Pre-1992	Pre-1992
Showerhead flow rate (gpm)				
<= 2.50	69.3%	67.0%	59.8%	65.1%
2.51-2.75	3.6%	4.7%	2.7%	4.3%
2.76-3.00	9.2%	9.8%	10.7%	9.9%
3.00+	17.9%	18.5%	26.8%	20.7%
Percent of all showerheads that are leaking	5.8%	5.9%	6.9%	6.2%

LFSHs and the incidence of leaking devices among owner-occupied multi-family units (e.g. high-rise condominiums) should be similar to that of single-family units. Accordingly, the summary statistics shown for multi-family units may over-state the actual saturation of LFSHs among apartment units. If so, it may be most cost-effective for utilities to develop programs that specifically target apartments for future LFSH device retrofits.[62]

Reasons for Replacing Showerheads

Occupants of pre-1992 housing reportedly replaced more than half (55.3%) of all showerheads during the eight years between 1992 and 2000 (see Table 9). The replacement rate differed markedly between single-and multi-family units, however, as 63.5% of all showerheads were reportedly replaced during this period in single-family compared with only 37% in multi-family dwellings.

[62] Because of the way dwelling units were classified for purposes of the study, it is not easy to distinguish apartment units from other types of multi-family dwellings. However, an attempt will be made during future analysis to identify and estimate separately saturation rates for apartment units.

Table 9: Reasons for replacing showerheads

Reasons given by those reporting the showerhead had been replaced	Single-Family		Multi-Family	Full Sample
	Thru 2000	Pre-1992	Pre-1992	Pre-1992
Broken	35.6%	33.7%	47.9%	36.8%
Old	8.6%	14.3%	9.3%	8.1%
Remodeling	39.9%	35.9%	25.7%	38.8%
Save water	6.5%	7.9%	2.1%	6.6%
Utility program	6.6%	6.3%	5.7%	6.2%
Personal preference	1.8%	1.0%	5.3%	2.0%
Other	1.0%	0.9%	3.9%	1.6%

The vast majority (83.7%) of showerheads were reportedly replaced because they were broken, old or during remodeling. Despite the large number of low-flow devices distributed by utility-funded programs during the past 12 years, only 6.2% of study households said they had replaced an old showerhead with a utility-provided low-flow device.[63] Approximately the same share of all households reported replacing showerheads to save water, although single-family households were more likely than occupants of multi-family dwellings to mention conservation as a reason for replacing showerheads.

2. Toilets

Flush Volumes and Leaks

Flush volumes for all inspected toilets are summarized in Table 10. As can be seen, the saturation of ULFTs is noticeably higher for single-family dwellings. The large percentage of 3.5 gpf toilets is also noteworthy.

At first glance, these findings would seem to raise important questions about the cost-effectiveness of future toilet retrofit programs assuming the typical retrofit would save only two gallons per flush. But this initial impression must be tempered by anecdotal evidence from the field suggesting that manufacture-rated 3.5 gpf toilets often flush far more than 3.5 gallons in actual use. Unfortunately, since most toilets were classified by inspectors based on their physical and operating characteristics, the actual volume of water flushed by the typical 3.5 gpf device is unknown. Consequently, the cost-effectiveness of replacing 3.5 gpf cannot be determined based on the available data. This underscores the need for additional research to provide good estimates of the amount of water saved by retrofitting the typical 3.5 gpf unit.

[63] It is important to note that the question about reasons for replacing showerheads was only asked of those who reported replacing one or more showerheads while occupying the dwelling. Given the large number of low-flow showerheads distributed by utilities during the early 1990s and the fact that the average tenure for occupants of pre-1992 study households was 8.5 years (see Table 7), the share of LFSH retrofits attributable to utility-funded programs is likely to be higher than 6.2%. In other words, a utility-provided device may have been installed before the current occupant moved into the unit.

Table 10: Toilet flush volumes and leaks

Variable	Single-Family		Multi-Family	Full Sample
	Pre-1994	Pre-1992	Pre-1992	Pre-1992
Toilet flush volume (gallons)				
1.60	48.6%	48.6%	37.6%	45.8%
3.50	35.3%	35.4%	47.6%	38.5%
5.00	15.6%	15.3%	13.9%	14.9%
7.00+	0.5%	0.6%	1.0%	0.7%
Percent of all toilets that are leaking	3.0%	3.1%	5.1%	3.6%

The estimates in Table 10 have potentially important implications for the design and the cost-effectiveness of future toilet retrofit programs. In contrast to showerheads, there is no reason to assume that the saturation of ULFTs, or the savings potential available from replacing toilets, is geographically uniform. To the contrary. Field experience and economics both suggest that the remaining stock of high flush volume toilets tends to be disproportionately concentrated in lower income neighborhoods and in renter-occupied units. While pursuing this conjecture is beyond the scope of the current analysis, it deserves close scrutiny in future research. The practical implication may be that future ULFT retrofit programs will need to make a greater effort to identify and selectively target neighborhoods and dwellings offering the greatest potential water savings per dollar invested. It is quite likely that these programs will need to be designed and operated quite differently from the mass distribution and rebate programs that have been a staple of utility-assisted residential ULFT retrofit activity.

Reasons for Replacing Toilets

For the sample of pre-1992 dwellings as a whole, 37.3% of all toilets were reportedly replaced between 1992 and 2000. As was true for showerheads, there are notable differences between occupants of single-family units and those living in multi-family units. Whereas half (50.3%) of all toilets in single-family units were reportedly replaced during this period, only 28.1% of the toilets in multi-family units were said to have been replaced.

Table 11 lists the reasons given for replacing toilets by households saying that that a toilet had been replaced while they were living in the unit. As with showerheads, most devices were reportedly replaced because they were broken or because a bathroom was being remodeled. In contrast to the case of showerheads, however, a significant proportion of toilets—one quarter of the full sample of pre-1992 units—were reportedly replaced as the direct result of a utility-funded program.

Table 11: Reasons for replacing toilets

Reasons given by those reporting the toilet had been replaced	Single-Family		Multi-Family	Full Sample
	Thru 2000	Pre-1992	Pre-1992	Pre-1992
Broken	31.1%	29.7%	42.3%	31.9%
Old	2.0%	2.1%	12.3%	3.8%
Remodeling	37.0%	37.8%	23.3%	35.4%
Save water	2.0%	2.1%	2.5%	2.2%
Utility program	27.9%	27.8%	12.8%	25.3%
Personal preference	0.0%	0.4%	0.0%	0.3%
Other	0.0%	0.0%	6.8%	1.1%

The pattern of differences between single- and multi-family units regarding showerhead retrofits appears to largely repeat itself here. Toilets in single-family households were most likely to have been replaced during a bathroom remodeling project, whereas those in multi-family units tended to be replaced out of necessity—that is, because they were broken or old. Single-family households were also more than twice as likely than residents of multi-family units to report replacing their toilet through a utility-funded program. In part, this probably reflects the fact that renters are likely to feel the need to obtain permission from the property manager or building owner to replace their toilets, a prospect that may effectively deter their inclination to participate in utility-sponsored programs. Additionally, since most renters do not receive water bills, and because may utilities use billing inserts to publicize their retrofit programs, renters may be less likely than single-family households to know about the existence of these types of programs. Once again, this suggests the possible need to tailor programs and to selectively target specific segments of the multi-family housing market for device retrofit programs. Retail utilities in Southern California have historically recognized and attempted to address this need through a combination of creative outreach, community partnering and device distribution strategies. But it may be time to revisit and to update these targeting strategies.

3. Program Freeriders

As previously mentioned, California state law has restricted the sale and installation of showerheads, toilets and faucet aerators to water-conserving models since 1992. Rather than passively wait for all water-wasting toilets and showerheads to be replaced through the process of *natural replacement*, utilities have continued to fund device retrofit programs in an attempt to motivate households to replace water wasting but still functioning devices with their water-efficient alternatives earlier than would normally occur. In this way, the utility attempts to accelerate the date when conservation savings being to accrue. If the cost savings realized by the utility as the result of these accelerated retrofits exceed program costs, the retrofit is cost-effective and both the utility and its customers experience a net benefit.

Some people who receive free or subsidized water conserving toilets and showerheads through utility-funded retrofit-programs would have replaced the device(s) at their own expense had the

utility's program not existed. This might happen, for example, when the device has failed or is replaced during a household remodeling project. These types of program participants are referred to as program *freeriders* because they opportunistically use program resources to subsidize a purchase that would otherwise be fully paid for at private expense. Because the resulting retrofit does not accelerate water savings beyond what would have occurred without the program, the utility incurs a cost without a corresponding increase in water savings. In this way, program freeriders directly reduce the cost-effectiveness of utility funded retrofit programs. Unfortunately, because good empirical estimates of water utility program freeridership are virtually nonexistent, it is difficult to adjust standard program cost-effectiveness estimates to properly account for the effects of freeriders.[64] The *Orange County Saturation Study* provided the opportunity to develop some preliminary estimates of program freeridership.

During the on-site inspection participants who reported replacing a showerhead or toilet were asked if they had obtained the device(s) privately or through a utility-sponsored program. Those saying they had obtained the device(s) through a utility program were then asked if they would have replaced the device(s) at their own expense had the public program not been available. Given the hypothetical nature of the questions asked and the problematic reliability and validity of respondent self-reports, discussed earlier, extreme caution needs to be used when interpreting the answers given to these questions. Indeed, the information presented here is most appropriately viewed as illustrative, not as an empirically sound basis for developing meaningful estimates of program freerider rates.

Table 12 shows the total number of toilets in households comprising the county sample that had reportedly been replaced since 1992—the year in which legislation restricting the installation of showerheads and toilets to water conserving models took effect. Among the 67 toilets said to have been replaced through a utility-funded program at single-family units, approximately 29% (20/67) would reportedly have been replaced anyway at private expense if the utility program had not existed. The corresponding estimate of program freeriders among occupants of multi-family units is 18% (1/6).

As noted earlier, the estimates of program freerider-ship developed here must be interpreted and used with extreme caution. Aside from the problematic reliability of respondent self-reports, the number of households used in developing the estimates is uncomfortably small. Additionally, the validity of answers given to questions like those used in the survey is an open issue. Generally speaking, credible estimates of program freeriders requires a different research design, more extensive questions, and more sophisticated analytic methods than those used in this study.

[64] The California Urban Water Conservation Council's forthcoming study of the incidence of freeriders among ULFT retrofit program participants should help fill this void.

Table 12: Toilet retrofits funded at private and at public expense

Variable	Single-Family		Multi-Family	
Toilets replaced at private expense	141		24	
Toilets replaced through a utility-sponsored program	67		6	
Number that would otherwise have been replace at private expense		47		5
Number. that would not have otherwise been replaced		20		1
Toilets replaced, source unknown	16		9	
Total number of toilets replaced	224			

D. Other Indoor Plumbing Devices

This section contains estimates of the prevalence and use of dishwashers and clothes washers. These estimates should be useful to conservation planners and those who use end-use models to forecast future water requirements.

1. Dishwashers

Table 13: Dishwasher saturation and frequency of use

Variable	Single-Family		Multi-Family	Full Sample
	Thru 2000	Pre-1992	Pre-1992	Pre-1992
Percent households with a dishwasher	83.0%	82.1%	65.8%	76.3%
Average dishwasher runs per week	2.9	2.9	2.4	2.7

2. Clothes Washing Machines

Table 14: Clothes washing machine saturation and frequency of use

Variable	Single-Family		Multi-Family	Full Sample
	Thru 2000	Pre-1992	Pre-1992	Pre-1992
Percent households with an in-unit machine	96.5%	96.6%	25.6%	71.5%
Percent households without an in-unit machine but with access to a machine on the property	2.7%	2.5%	69.7%	26.3%
Percent households without on-site access to a machine	0.8%	0.8%	4.6%	2.2%
Totals	100.0%	100.0%	100.0%	100.0%
Average washing machines runs per week (units with in-unit machines only)	6.8	6.8	5.2.	6.6

Note: Average washing runs per week are estimated only for households with in-unit machines.

VI. CONCLUSIONS

A. Study Objectives

The *Orange County Saturation Study* had two main objectives. The first was to test the hypothesis that data from a county-level sample of households conducted in Southern California will produce good estimates of the saturation of low-flow showerheads (LFSHs) among households served by the separate retail water agencies operating with the county. The second was to test the proposition that the 75% showerhead saturation level required by CUWCC's BMP 2 has been achieved.

These primary study objectives fostered several subsidiary objectives. The first was to develop a research design capable of generating the data needed to empirically test the two study hypotheses. In the process, it was recognized that valuable information could also be collected about the type and use of other residential indoor water-using devices at a relatively small increase in total study costs. Given the importance of this type of information for designing and administering cost-effective residential water use efficiency programs, the scope of the study was expanded to provide the data needed to develop a rudimentary profile of indoor water use. A grant from the USBR also allowed information to be collected about household participation in utility-sponsored plumbing retrofit programs and tentative estimates of the incidence of program freeriders.

B. Study Findings

1. Estimating Low-Flow Showerhead Saturation

Testing the Hypothesis of Geographically Uniform Low-Flow Showerhead Saturation Rates

Data from two independent samples were used to test the proposition that the results of county-level surveys can be used to estimate the saturation of LFSHs among household served by the separate retail utilities within the county. The first was a representative sample of all pre-1992 dwellings in Orange County. The second consisted of a sample drawn from two contiguous cities in Orange County, each of which is served by a separate municipal water utility.

Separate saturation rates were estimated and compared for single-family households in the city and the county samples. The same was done for multi-family dwellings. In both cases, the estimates obtained from the city and the county samples were statistically equivalent. This result, which holds true for both single- and multi-family dwellings, provides strong empirical support for the hypothesis of geographically uniform saturation of LFSHs throughout Orange County.

These findings supports the proposition that, at least in Southern California, a representative sample of county households can provide good estimates of the existing saturation rate for geographic sub-units within the county—such as an individual retail utility's service area.

Testing the Hypothesis that the 75% Saturation Threshold has been Achieved

This hypothesis was initially tested following the statistical criteria prescribed in CUWCC's *MOU*. This was done by calculating separate 95% confidence intervals for randomly drawn subsets of single- and multi-family dwellings from the county sample.[65] The upper bound of the 95% confidence interval for single-family units just barely captured the 75% threshold while the confidence interval for multi-family units did not. Technically, this means that, as of the winter of 2000, the 75% saturation requirement had been satisfied among single-family household in Orange County. But this interpretation, while statistically defensible, is tenuous given more compelling evidence suggesting that the true LFSH saturation rate for single-family households during the winter of 2000 was probably much closer to 67% than it is to 75%. For multi-family units the actual saturation rate most likely ranged somewhere between 53% and 66%.[66]

2. The Saturation of Ultra-Low Flush Toilets

The county sample of Orange County households indicated that roughly 48% of all toilets in single-family housing units were ultra-low flush (ULFT) models. The saturation of ULFTs among multi-family dwellings, comprised largely of apartments, lagged behind at about 38%. The remaining toilet population consisted predominately of 3.5 gpf models in both single- and multi-family dwellings, a finding that may indicate the need to re-examine current water-savings estimates attributed to ULFT retrofits.[67]

3. Noteworthy Differences Between Single-Family and Multi-Family Dwellings

As is true for the saturation of ULFTs, multi-family dwellings lag behind single-family units when it comes to LFSH saturation. Part of this difference is explained by the reasons given for replacing these types of devices. Occupants of single-family dwellings are much more likely to report having replaced showerheads and toilets during a remodeling project or, in the case of toilets, because they obtained a ULFT through a utility-sponsored program. Occupants of multi-family dwellings, on the other hand, were more likely to report that their showerheads and toilets had been replaced only because the old ones had failed.

The lower saturation of water-efficient devices among multi-family dwellings, and the fact that they are less likely to report participating in utility-sponsored retrofit programs, suggests several possibilities. They may be less aware of the existence of utility-sponsored programs. They may also face greater barriers to program participation than do occupants of single-family units.

[65] Random subsets were drawn to obtain sample sizes equal to those required by the *MOU*'s statistical test criteria.
[66] See Chapter IV.
[67] See Section V.C.2 for elaboration.

While conservation planners have long recognized these possibilities and attempted to overcome them through creative outreach and targeted program strategies, it may be time to revisit this issue.

4. The Reliability of Respondent Self-Reports

Because of concerns about the questionable accuracy of self-reported information about the presence of water-efficient devices in their homes, trained inspectors visited each study household to precisely measure showerhead flow rates and toilet flush volumes. As discussed in Section III.E.3, basing saturation estimates on the information provided by household occupants during the telephone interview would have underestimated the true saturation of LFSHs and ULFTs by about 34% and 42% respectively. Errors of this magnitude raise serious concerns about the validity of saturation estimates based on telephone surveys alone.

C. Qualifications

As discussed in Chapter III and in several of the appendices, the geographic distribution of pre-1992 households comprising the county sample closely parallels the distribution of the total pre-1992 housing stock in Orange County. Additionally, there is no identifiable difference between those households who participated in the telephone screening and enumeration survey but subsequently declined to participate in the on-site inspection phase of the study and those who did participate. Nonetheless, appropriate caution is necessary generalizing these findings to areas outside of Southern California.

It is also important to recognize that most of the analysis described in this report is, by design, limited to the existing stock of pre-1992 units in Orange County. This is because the retrofit activity of primary interest to the study only involves this segment of the housing stock. At the same time, however, the original intention was to also generate a representative sample of the entire year 2000 housing stock. The purpose was to provide a representative profile of the type, incidence and use of those indoor water-using devices that are not currently subject to water-efficiency standard and regulations. As discussed in Chapter III, however, for unknown reasons, the supplementary sample of post-1991 units did not include any multi-family units. While this turns out to be of little practical consequence, it is a fact that needs to be kept in mind, especially where the results discussed in Section V.D of this report are concerned.

APPENDIX A: SAMPLE SIZE CALCULATIONS

Under the California Urban Water Conservation Council's revised *Memorandum of Understanding Regarding Urban Water Conservation in California* (referred to hereafter as the *urban MOU*), an agency can be certified as having fully implemented *Best Management Practice #2* (BMP 2) if it

> can demonstrate through customer surveys with 95% statistical confidence and a ±10% error that 75% of single-family residences and 75% of multi-family units constructed prior to 1992 are fitted with low-flow showerheads.[68]

Standard statistical calculations show that if "true" LFSH saturation is at 75% for the population as a whole, a random sample of 75 showerheads[69] is sufficient for estimating average saturation with a 95% confidence interval not exceeding ±10 percentage points. This result is obtained by applying a normal approximation to the true underlying binomial distribution, a strategy that is accepted practice in survey sampling, especially when design effects due to intra-cluster correlation and sample weighting need to be taken into account.

Although statistical theory provides straightforward guidance about the number of showerheads that need to be sampled to satisfy the prescribed statistical criteria, determining the number of separate dwelling units that need to be inspected is more complicated.

Table A1 shows the distribution of full bathrooms, by housing type, for the pre-1992 housing stock in the Anaheim-Santa Ana Metropolitan Statistical Area based on the 1995 Annual Housing Survey.[70] The estimates presented in this table show, for example, that more than three-quarters of all multi-family units have one full bathroom and few, if any, have more than two. In contrast, slightly less than one-quarter of all single-family units have only one full bathroom, while nearly two-fifths have more than two full bathrooms.

A1: Distribution of full bathrooms, by housing type

	Distribution of Pre-1992 Housing Stock	
Number of full bathrooms	**Single-Family**	**Multi-Family**
One	23.2	78.3
Two	60.8	21.7
Three	13.5	≈0
Four or more	2.6	≈0
Total	100.0	100.0

SOURCE: American Housing Survey, 1995.

[68] CUWCC *MOU*, as amended September 21, 2000, p. 21.
[69] It is important to note here that the appropriate unit of analysis in this case is the showerhead, not the dwelling. Accordingly, fewer than 75 households may be required to yield a sample of 75 showerheads. This fact is discussed at greater length in the text.
[70] Information about the Annual Housing Survey can be found at www.census.gov/hhes/www/housing/ahs.

Based on the distributions shown in Table A1, a random sample of households should, on average, yield 1.95 and 1.21 showerheads per single-family (SF) and multi-family (MF) dwelling unit, respectively. From this, one might conclude that fewer than 75 dwellings of each housing type need to be sampled, since the sample only requires 75 SF and 75 MF showerheads. But this conclusion would be wrong because it ignores the effects of intra-cluster correlation and sample weighting.

Intra-Cluster Correlation

One source of sample design effects that needs to be considered is the clustering of low-flow showerheads within a dwelling unit. To the extent that either all or none of the showerheads in a unit are low-flow models, a strong clustering effect results. This clustering effect degrades the effective sample size because the status of one showerhead perfectly predicts the status of all showerheads in a dwelling unit. In other words, once the status of the first showerhead is determined, no new information is added by determining the status of all remaining showerheads. In such an extreme case, randomly selecting one showerhead per sampled dwelling is enough. It therefore follows that obtaining a sample of 75 showerheads would still require 75 dwelling units. This extreme case is discussed only for illustration. In practice, the clustering effect depends upon both the total number of showerheads per dwelling unit and the proportion of low-flow showerheads within the typical dwelling.

To demonstrate the effects of inter-cluster correlation (δ), Table A2 presents four scenarios for the population of SF dwelling units having two showerheads. In each case, the average LFSH saturation rate remains constant at 75%. Scenario 4 assumes that either both, or neither of the showerheads, are low-flow (δ=1.00). The other scenarios vary the proportion of dwellings with 0, 1, and 2 low-flow showerheads to simulate lower levels of inter-cluster correlation.

A2: The effect of inter-cluster correlation on effective sample size

	Scenario			
	1	**2**	**3**	**4**
Percent of SF dwellings with 0 LFSH	6.25%	12.50%	18.75%	25.00%
Percent of SF dwellings with 1 LFSH	37.50%	25.00%	12.50%	0.00%
Percent of SF dwellings with 2 LFSH	56.25%	62.50%	68.75%	75.00%
LFSH penetration rate across dwellings	75.00%	75.00%	75.00%	75.00%
Intra-cluster correlation (δ)	0.00	0.33	0.66	1.00
DEFF (design effect)	1.00	1.33	1.66	2.00
Actual showerheads per dwelling	2.00	2.00	2.00	2.00
Effective showerheads per dwelling	2.00	1.50	1.20	1.00

As Table A2 shows, the size of the design effect directly corresponds to the extent of degradation in effective sample size because of inter-cluster correlation. Under Scenario 1, for example, the design effect equals 1.00, meaning no degradation has occurred. This is why effective number of showerheads per dwelling remains unchanged at two. In contrast, for Scenario 4 the design effect equals 2.00, meaning inter-cluster correlation has caused the effective sample size to be reduced by half.

Without strong reasons to favor one scenario over the other, the somewhat conservative Scenario 3 was selected for planning purposes. Assuming the design effect is also 1.66 for SF dwellings with three and four showerheads (for SF dwellings with one showerhead, design effect is 1.00 by definition), it was estimated that approximately 60 randomly selected SF dwelling units would be required to yield an effective sample size of 75 showerheads. In contrast, since most MF units have only one showerhead, 75 MF dwellings seemed a reasonable estimate of the required sample size for these types of units.

Weighting

Design effects also result from the use of weights intended to correct for commonly occurring discrepancies between the composition of the sample and the population from which it is drawn. For simplicity, assume there are two classes of households—frequent responders and infrequent responders. Further assume that the infrequent responders make up 30% of the sample but a larger share of the full population. In this case, data for the infrequent responders would need to be up-weighted while data for the frequent responders would need to be down-weighted to achieve a proper correspondence between the characteristics of the sample and those of the general population. The more the ratio of these two weights deviates from 1.00, the more the discrepancy between the sample and parent populations.[71] Table A3 presents several examples.

[71] Sample weights larger than 5 or so become grounds for questioning the quality of the sample. They also lead to practical concerns that the use of such large sample weights may further distort rather than help to improve the quality of sample-based estimates.

A3: Design effects associated with disproportionate response rates—hypothetical example

	Scenario			
	A	**B**	**C**	**D**
Proportion of infrequent responders in sample	30.00%	30.00%	30.00%	30.00%
Ratio of weights applied to frequent and infrequent responders	1.50	2.00	3.00	5.00
Design effect	1.04	1.12	1.33	1.69

Under the worst-case scenario (Scenario D), the ratio of the two group weights is assumed to be 5.00, which yields a design effect of 1.69. In other words, application of these weights, although necessary for removing the effects of disproportionate response rates, simultaneously reduces the effective sample size by a factor of 1.69 under Scenario D.

For planning purposes, Scenario C in Table A3 was used. Thus, taking into account the design effect caused by weighting raises the target sample size for MF dwelling units from 75 to roughly 100 (75×1.33). Similarly, the size of the target sample for SF dwellings increases from 60 to roughly 80 (60×1.33). Ultimately, in an effort to ensure that the resulting sample would satisfy the prescriptive statistical requirements of BMP 2, the final sampling targets were set at 100 dwelling units for each housing type.

Field Results

In practice, the actual size of sample design effects turned out to be less than anticipated. This happened because the intra-cluster correlation was lower than expected and because of the good geographic balance achieved by the county sample. Combined, these results had the desirable effect of eliminating the need for heavy weighting. Consider, for example, the issue of intra-cluster correlation.

Table A4 shows the actual incidence of low-flow showerheads among SF dwellings in the county sample with two and three showerhead. For example, in the 226 SF dwellings with two showerheads, both turned out to be low-flow in 54.4% of the cases, one in 30.1% of the cases, and none in 15.5% of the cases. This distribution is much closer to Scenario 2 in Table A2 than it is to Scenario 3, the scenario used in calculating sample size requirements. The distribution is even more uniform among SF units in the county sample with three-showerheads per dwelling. Intra-cluster correlation for sampled units with two and three showerheads per unit works out to 0.29 and 0.09 respectively. With respect to the effects of weighting, the design effects calculated for the sample data are actually closer to Scenario A in Table A3 than to Scenario C, the one used for planning purposes. Consequently, weighting reduced the effective sample size less than was originally anticipated.

A4: Incidence of low-flow showerheads for selected single-family dwellings (county sample)

	Breakdown by number of low-flow showerheads found				
	0	**1**	**2**	**3**	**Total**
SF dwellings with two total showerheads	35 (15.5%)	68 (30.1%)	123 (54.4%)		226 (100.0%)
SF dwellings with three total showerheads	3 (5.8%)	15 (28.8%)	18 (34.6%)	16 (30.8%)	52 (100.0%)

To assist in the planning of future studies, Table A5 summarizes total design effect estimates for the county and the city samples by dwelling type. This table gives the mean LFSH saturation rates and both the uncorrected and corrected standard errors associated with these estimates. The squared ratio of uncorrected to corrected standard errors represents the total design effect, accounting for both the effects of intra-cluster correlation and sample weighting. There is, of course, no guarantee that exactly the same design effects will occur in other saturation studies. Housing stock characteristics may be different, and tardy implementation may necessitate the use of larger weights. Therefore, the design effect estimates shown in Table A5 should be used with appropriate caution when planning future studies.

Table A5: Design effects and corresponding corrections for the city and county samples

Variable	**County Sample**		**City Sample**	
	SF	**MF**	**SF**	**MF**
No. of inspected dwelling units (pre-1992)	358	196	85	93
No. of inspected showerheads	703	258	145	119
Estimated LFSH saturation (mean)	66.92%	59.76%	66.21%	54.62%
Corrected standard error	1.98%	3.31%	4.20%	4.85%
Uncorrected standard error	1.77%	3.05%	3.93%	4.56%
DEFF (design effect)	1.25	1.18	1.14	1.13

APPENDIX B: TELEPHONE SCREENING AND ENUMERATION SURVEY

Telephone-based survey research has evolved into a sophisticated and highly specialized activity over the past 30 years. Some firms, for example, specialize in developing nationwide databases for use in conducting random digit-dialed (RDD) telephone surveys. These databases usually contain area codes, telephone prefixes, and associated geographic information like zip codes obtained from telephone companies and other industry sources. The telephone prefixes are generally screened to eliminate non-assigned numbers, fax machines, pagers and other non-voice numbers. Since telephone companies tend to assign different prefixes to residential and non-residential numbers, these databases can be further filtered to produce a sample of telephone number prefixes that increase the likelihood that a randomly dialed number will reach a working residential telephone number. Farrand Research, the company responsible for coordinating the telephone screening survey for the *Orange County Saturation Study* purchased two such lists from Scientific Telephone Samples. One consisted of a random sample of telephone numbers for all of Orange County. The second consisted of a random sample of numbers for the two cities of Buena Park and Fullerton.[72]

The actual telephone interviews were conducted by Western Watts using a computer-assisted telephone interviewing (CATI) system. The system randomly selected and automatically dialed individual telephone numbers from the purchased list of telephone numbers. If a number corresponded to a study-eligible residence and the respondent agreed to participate, the survey was administered and answers were entered, on a real-time basis, into the CATI system's database. If the telephone contact resulted in a *terminal outcome* (e.g., a non-voice or non-residential number was reached; a household located outside the study area was dialed; a respondent declined to participate in the telephone interview; or a completed interview was obtained), the telephone number was deleted from the sampling frame. If the outcome was not terminal (e.g., no answer, answering machine intercept, the respondent asked to be called at a later time), the CATI system automatically scheduled the number for redialing at a different day and time. Each randomly selected telephone number was called up to a maximum of seven times or until a terminal outcome was obtained.

The telephone survey was administered in both English and Spanish. Consideration was initially given to conducting interviews in other languages, but the cost of this enhancement could not be justified. According to the 1990 Census, for example, while Vietnamese is the third most frequently spoken language in Orange County, only about 3% of the total Orange County population speak this language.

[72] Because telephone area codes and prefixes are seldom coterminous with city or county boundaries, care needs to be taken to ensure that the sample accounts for all telephone numbers within these geographic areas. This means that the sampling frame contained all telephone numbers belonging to area codes or prefixes that may be bisected by county or city boundaries. To establish the eligibility of randomly contacted households, telephone interviewers had to ask if the household was located within Orange County (county sample) or in Buena Park or Fullerton (city sample). Households located outside these geographic boundaries were flagged as not eligible for the study and the telephone contact was terminated.

Overall, 10% of Orange County's 1990 population reported not being fluent in English. But the proportion of households without a single English-speaking occupant is likely to be much smaller. In light of the extremely small probability of reaching a household without an English- or Spanish- speaking member, language barriers should not be a significant source of bias in the final samples.

For telephone numbers with repeated answering machine intercepts, interviewers were instructed to leave a brief message on the sixth and seventh call describing the purpose of the call. This strategy was adopted to balance two considerations—not annoying customers with repeated messages while attempting to establish contact with those households who routinely use their answering machines to screen telephone calls.

The CATI system was programmed to partition the full set of randomly generated telephone numbers into smaller blocks, or replicates. All numbers belonging to a replicate were dialed a maximum of seven times or until a designated terminal disposition code was obtained, whichever came first, before calls were begun to numbers comprising the next replicate. This strategy is very important for obtaining a representative sample of households. Simply proceeding sequentially through a long list of telephone numbers, even a randomly ordered list, does not provide adequate assurance that the resulting sample will be truly representative. Early calls are more likely, for example, to reach the retired, the disabled, and the unemployed because these groups have the highest probability of being at home at any given moment. It is the repeated callbacks, varied by time of day and day of week, that bring the more difficult-to-reach households into the sample, thereby increasing the representativeness of the final sample. While making more than seven attempts to contact each sampled telephone number would, in principle, further enhance sample representativeness, this consideration had to be balanced against cost. The existing literature on telephone surveying suggests that the gains from dialing each number more than seven to nine times are very small. Market-research surveys usually set this threshold between three and five calls.

Care was given to crafting an opening script designed to maximize participation in the telephone screening interview. Each eligible respondent was asked several questions about the physical characteristics of their dwelling and about the number of bathrooms, toilets and showers in their residence. Information about structure type (single-family, multi-family, mobile home) and age was necessary to satisfy the targets set for each sample stratum. For the city sample it was also necessary to screen for age of structure, given the study's focus on the saturation of low-flow showerheads in pre-1992 units. Appendix C contains a complete copy of the telephone survey instrument.

Study-eligible households identified during the telephone screening survey were invited to participate in the on-site inspection phase of the study. Again, a carefully crafted recruitment script was used to maximize participation. At this point, prospective participants were also told that, after completing the on-site inspection, their household would be entered in a lottery from which 12 winners would be drawn at random to receive a $500 reward gift certificate. The main purpose of this incentive was to minimize attrition among those who initially agreed to participate in the on-site inspection

It was expected that some households would renege on their initial stated willingness to participate in the on-site inspections between the time of the telephone screening survey and the time they were contacted to schedule the on-site inspection. It was conservatively estimated that the attrition rate could run as high as 25%. Therefore, 1000 willing, study-eligible participants were recruited through the telephone screening survey in an effort to assure that 800 completed field inspections would be obtained.

For the city sample, respondent-reported information about dwelling type and construction date was continually monitored in an effort to assure that 100 SF and 100 MF completed field inspections would be obtained. For the county sample, however, screening for dwelling type was initially not undertaken under the expectation that the random-digit-dialing survey procedure alone would produce a mix of SF and MF households proportionate to their share of the total Orange County housing stock. As it turned out, SF dwellers were being recruited at a somewhat higher than expected rate, so midway during the telephone surveying, screening for dwelling type was implemented for the county sample. Why this over-recruitment of SF dwellings occurred remains unclear. We surmise that since renters (who have weak conservation incentives) predominate in the MF housing sector, fewer MF households were completing the telephone interviews. We do know that among the SF and MF households that completed telephone interviews, roughly the same proportion in each group also agreed to on-site inspections. Therefore, the over-representation of SF dwelling units was not the result of a higher on-site inspection agreement rate among occupants of SF dwellings.

By the time the telephone survey budget was exhausted, a total of 1,045 eligible households had been identified that were also willing to participate in the on-site inspection phase of the study. Of these, 72 required that additional authorization be obtained from someone other than the person who completed the telephone screening survey. Among the latter group, 36 households eventually participated in on-site inspections. Overall, the 1,045 eligible households yielded 762 completed inspections—38 short of the study target of 800.

Call Statistics

Call disposition statistics provide important data for costing similar surveys in the future. A total of 43,239 telephone numbers were dialed a total of 167,952 times until a terminal disposition code was obtained, or until the seven-call limit was exceeded, whichever occurred first.

Table B1 tabulates the final status of each number dialed. Slightly over a quarter (28%) of the calls never reached a terminal disposition status (e.g., no answer, busy signal), a rate that is generally consistent with the published literature on telephone surveys. Almost 13% were disconnected numbers. Many other numbers were either ineligible (e.g., a business or a residence not located in Orange County) or non-voice numbers (e.g., pagers and fax machines). Overall, it is estimated that the dialed numbers included a total of at least 18,120 eligible residential numbers, ignoring for the moment eligible residences that might be included in the "no answer" and "busy signal" categories.

Table B1: Frequency of each dialed telephone number's final disposition

Disposition Code	Description	Frequency	Percentage	Likely, Eligible, Residential Numbers
01	No answer	11,128	25.74%	
02	Busy signal	926	2.14%	
03	Answering machine	1,771	4.10%	1,771
04	Not available	563	1.30%	563
05	Callback	353	0.82%	353
06	Disconnect	5,593	12.94%	
07	Business	3,883	8.98%	
09	Language barrier (not Spanish)	770	1.78%	770
10	Respondent terminate	1,115	2.58%	1,115
13	Refusal	7,818	18.08%	7,818
14	Completed seven calls with no answer	32	0.07%	32
15	Wrong telephone number reached	473	1.09%	
16	Not a residence	460	1.06%	
17	Not a residence in Orange County	381	0.88%	
18	Respondent not a resident of dialed home and no resident available	139	0.32%	139
19	Respondent not 18+ years old, no 18+ person available	201	0.46%	201
20	Completed telephone interview	2,802	6.48%	2,802
21	Refused to answer bill payer question	44	0.10%	44
22	Does not live in one of the listed O.C. cities (county sample) or Buena Park and Fullerton (city sample)	1,735	4.01%	
23	Refused zip code information	35	0.08%	35
24	Answered "don't know," "refused," or "other" for dwelling type	21	0.05%	
25	Dwelling constructed after 1994 (city sample only)	233	0.54%	
26	Residence is not in Orange County (Note: respondent may identify a community which is not officially a city/town, but is commonly used to identify the locale; see Code 22 above)	18	0.04%	18
30	Spanish language barrier (routed to Spanish interviewer)	304	0.70%	304
31	Pager	22	0.05%	
32	Fax machine	264	0.61%	
33	Terminated because quota for that type of household, either single or multi-family, already met (city sample only)	2,155	4.98%	2,155
	Total	**43,239**	**100.00%**	**18,120**

Table B2, on the following page, summarizes the outcomes of the 167,952 separate telephone calls made to identify and recruit study participants.

Table B2: Overall call disposition summary, rates and ratios

Call Statistics Summary	Numbers and Percentages
Overall Summary	
Total calls made	167,952
Total numbers dialed	43,239
Total eligible numbers found (best estimate)	18,120
Number that completed telephone surveys	2,802
Number that agreed to on-site inspections	1,045
Rates	
Incidence rate	41.9%
Telephone survey response rate	15.5%
On-site inspection agreement rate	37.3%
Ratios	
Average calls per dialed number	3.9
Average calls per completed interview	59.9
Average calls per inspection-ready household	160.7

Some of the more important results shown in this table include the following:

- The incidence rate (probability of dialing an eligible residence from the random list) works out to roughly 42% (18,120÷43,239); the telephone survey response rate was about 15% (2,802÷18,120).

- The on-site inspection agreement rate was approximately 37% (1,045÷2,802).

- On average, it took almost 161 calls to identify one eligible household willing to participate in the on-site inspections (excluding what it further took the field survey team to schedule appointments).

Except for on-site participation rate, which exceeded planning expectations by more than a factor of two, all the other ratios fell within anticipated ranges.

Appendix C: Telephone Survey Instrument

Orange County Saturation Telephone Screening Survey

Hello, this is _____ calling on behalf of the Metropolitan Water Districts of Orange County and Southern California. I would like to ask you a few questions about water usage in your home. Your answers will help us decide how to best meet Orange County's future water needs. Please be assured we are not selling anything and all your answers will remain confidential. This call will only take about five minutes. When we are finished, you may be invited to participate in the second part of the study which would take place at a later date.

SAMPLE REPLICATE NUMBER:___ LANGUAGE: SPANISH:___ ENGLISH:___

> **READ ONLY ON 5TH OR LATER ATTEMPT**: Hello, I'm calling on behalf of the Metropolitan Water Districts of Orange County and Southern California. We would like to ask you a few questions about water usage in your home. We will try again during the next several days. Thank you.

SCREENER = *Q. A through Q. M*

A. Would you have five minutes to answer these few questions now?
   Yes ___ Proceed to B.
   No ___ When may we call back? Day _____ Time _____
   Refused ___ (THANK AND TERMINATE)

If respondent objects, read: 60% of Southern California's water is imported, which makes it a critical resource. The purpose of this study is to help local water agencies plan for what they will need to do to meet Southern California's future water needs. Your answers to our questions will help us do that.

If respondent asks about sponsors, read: The Metropolitan Water District of Southern California and the Metropolitan Water District of Orange County, or MWDOC (interviewer: pronounced "mow-doc"), are public agencies who obtain water from outside our area and distribute it to local water agencies in the county, including the one that supplies your water. If you wish further information about these public agencies or about this study, you may call Ms.**XXXXXXXX** at the Municipal Water District of Orange County at **XXX-XXX-XXXX**.

If respondent asks about how his/her phone number was obtained, read: Your number was chosen at random from among all the residential phone numbers in your area.

If respondent asks about interviewer or call center, read: Farrand Research designed this survey; it is a research firm that specializes in studying the public's opinions. They are located in Long Beach, California. (*Only if necessary):* We are calling from a telephone facility located in central Utah.

B. May I verify that I have reached XXX-XXX-XXXX (phone number)?
Yes, correct phone number _____ - continue
No, incorrect phone number _____. I'm sorry, I've dialed the wrong telephone number. (THANK AND TERMINATE)
Ref ___ - (THANK AND TERMINATE)

C. Since your number was chosen at random, I need to start with a few questions about your household. Is this a residence?
Yes _____ - continue
No _____. We are only interviewing to residences for this survey. Thank you for your time. (THANK AND TERMINATE)
Ref ___ - (THANK AND TERMINATE)

D. Is this residence in Orange County?
Yes _____ - continue
No _____. We are only interviewing to Orange County residences for this survey. Thank you for your time. (THANK AND TERMINATE).
Ref ___ - (THANK AND TERMINATE)

E. Are you a resident in this home?
Yes _____ - continue
No _____ - ask to speak to a resident at least 18 years old. IF NOT AVAILABLE, CATI SCHEDULES CALLBACK.
Ref ___ - (THANK AND TERMINATE)

F. Are you 18 years old or older?
Yes ____ - continue
No ____ May I speak to someone in your household who is 18 or older? (IF NOT, CALL BACK AT ANOTHER TIME).
Ref ____ - (THANK AND TERMINATE)

G. Do you generally pay the bills for your household?
Yes ____ - continue
No ____ QG1. May I speak to the person who does generally pay the bills? (INTERVIEWER: If bill payer is not available, continue with next question (QH). If respondent asks why you want to speak with the bill payer, say "Some of the questions in this survey may be easier for the bill payer to answer, but it is not critical that we speak to that person".)
Ref ___ - (THANK AND TERMINATE)

H. In what town or city is your home located?
_____ (FOR "CITY" SAMPLE, IF NOT FULLERTON OR BUENA PARK, THANK AND TERMINATE.) (FOR "COUNTY" SAMPLE, RESPONDENT STILL QUALIFIES EVEN IF THEY REFUSE TO ANSWER).

01. Aliso Viejo	15. Irvine	29. Rancho Santa Margarita
02. Anaheim	16. Laguna Beach	30. Rossmoor
03. Brea	17. Laguna Hills	31. San Clemente
04. Buena Park	18. Laguna Niguel	32. San Juan Capistrano
05. Costa Mesa	19. Laguna Woods	33. Santa Ana
06. Coto de Caza	20. La Habra	34. Seal Beach
07. Cypress	21. Lake Forest	35. Stanton
08. Dana Point	22. La Palma	36. Trabuco Highlands
09. El Toro	23. Los Alamitos	37. Tustin
10. El Toro Station	24. Mission Viejo	38. Tustin Foothills
11. Fountain Valley	25. Newport Beach	39. Villa Park
12. Fullerton	26. Orange	40. Westminster
13. Garden Grove	27. Placentia	41. Yorba Linda
14. Huntington Beach	28. Portola Hills	98. Other <SPECIFY>
XX. Atwood	XX. Capistrano Beach	XX. Corona Del Mar
XX. Foothill Ranch	XX. Midway City	XX. Silverado
XX. Sunset Beach	XX. Surfside	XX. Trabuco Canyon

Other O.C. City (specify)
Other (non-O.C.)

Other _____. Is (CITY RESPONDENT MENTIONS) in Orange County? IF "YES" RECORD RESPONSE AND CONTINUE. IF NOT AN O.C. CITY, SAY: "I'm sorry we are only including residences in Orange County in this survey. Thank you for your time.

Refused _____. (THANK AND TERMINATE)

I. What is your zip code?

_____

Don't Know - Continue
Ref ___ - (THANK AND TERMINATE)

J. Which of the following best describes the building in which you live?" (READ LIST. MARK ONE ONLY)

A single-family house detached from any other house. _____
(Clarification, if necessary: a single house with open space on all sides; that is, not attached to another house.)

A single-family house attached to one or more other houses such as a condominium or townhouse _____

A building with between two and four units _____

A building with five or more units _____

A mobile home or trailer ____

Other ___ - DESCRIBE AND CATEGORIZE INTO ABOVE CATEGORIES IF POSSIBLE, OTHERWISE MARK AS OTHER. FOR BOTH COUNTY AND CITY SAMPLES – TERMINATE IF NOT ABLE TO UP-CODE INTO ABOVE CATEGORIES. SAY: "We are only interviewing specific types of households. Thank you for your time."

Don't know _____ We are only interviewing specific types of households. Thank you for your time.

Refused _____ (THANK AND TERMINATE)

K. Is there a separate residence located on a floor directly above or below yours?

Yes _____
No _____

Dwelling unit status is determined from combination of Q.J and Q.K. (i.e., if Yes to Q.K – then unit is multi-family. If no to Q.K – then unit is single family. See grid on next page for proper classification of dwelling types based on responses to Q.J and Q.K.
This will be an overquota terminate question for city sample after meeting quota for higher incidence dwelling type.

		SF-detached	**SF-attached**	**2-4 units**	**5+ units**	**DK**	**Ref**
Q.K Responses	**Yes**	Incompatible	Multi family	Multi family	Multi family	Not eligible	Not eligible
	No	Singe Family	Singe Family	Singe Family	Singe Family	Not eligible	Not eligible
	DK	Not eligible	Not eligible	Not eligible	Not eligible	Not eligible	Not eligible
	Ref	Not eligible	Not eligible	Not eligible	Not eligible	Not eligible	Not eligible

INTERVIEWER: IF "INCOMPATIBLE" RESPONSE ABOVE, ASK FOLLOW UP TO Q.K BELOW:
Just to help me clarify the type of building you live in, is there a separate residence located above or below the one you live in?

Yes _____ = Multi-family classification
No _____ = Single-family classification

K1. Dwelling unit type (as determined from QJ and QK):

Single family

Multi-family

L. How long have you lived at your current residence? (RECORD ANSWER BELOW)
___ years ___months - (IF MORE THAN 10 YEARS SKIP TO QUESTION #1)
- (IF LESS THAN 10 YEARS, CONTINUE)

Don't know _____
Refused _____

M. Was the building in which you live built before 1992?
Yes _____ - continue

No _____ - FOR "CITY" SAMPLE ONLY SAY: We're only including buildings or homes constructed prior to 1992 in this study. Thank you for your time.
FOR "COUNTY" SAMPLE, CONTINUE TO Q.1

Don't know _____ SAME INSTRUCTIONS AS FOR "NO" RESPONSE ABOVE.
Refused _____ SAME INSTRUCTIONS AS FOR "NO" RESPONSE ABOVE.

(For all respondents)

1. Does your house receive and pay a water bill?

   Yes _____

   No _____

   Don't know _____

   Refused _____

2. Is your residence … (READ LIST ONE BY ONE. CONTINUE READING LIST UNTIL RESPONDENT ANSWERS YES TO ONE OF THE ANSWERS BELOW

   Owned or being purchased by you or someone else living in your household?

   Yes___ skip to Q.3

   No___ continue

   Don't know _____ continue

   Ref___ continue

   Rented?

   Yes___

   No___

   Don't know _____

   Ref___

3. How many people live in your home on a permanent basis?
   Write in # _____
   Don't know _____
   Ref _____

Now I'd like to ask you a few questions about your house and water using appliances.

4. How many bathrooms do you have in your home? Record # _____ Continue
   Don't know _____ Continue
   Refused _____ Continue

5. How many of these bathrooms have showers in them?
   Record # _____ Continue
   Don't know _____ Continue
   Refused _____ Continue

(ASK Q.6A ONLY IF "8 OR MORE YEARS" IN Q. L)

6A. Have any of the showerheads in these bathrooms been replaced since 1992?

   Yes _____ - continue

   No _____ - Skip to Q. 8

   Don't know_____ - Skip to Q.8A

   Refused _____ - Skip to Q. 8A

(ASK Q.6B ONLY IF "LESS THAN 8 YEARS" IN Q. L OR, - IF "DON'T KNOW/REFUSED" IN Q.M.)

6B. Have any of the showerheads in these bathrooms been replaced while you've lived there?

Yes _____ - continue
No _____ - Skip to Q. 8
Don't know_____ - Skip to Q.8B
Refused _____ - Skip to Q. 8B

7. How many showerheads have been replaced during this time period?
Record # _____ Continue
Don't know _____ Continue
Refused _____ Continue

(ASK Q.8A ONLY IF "8 OR MORE YEARS" IN Q. L)

8A. Have any of the toilets been replaced since 1992?

Yes _____ continue
No _____ skip to Q.10
Don't know_____ - Skip to Q.10
Refused _____ - Skip to Q. 10

(ASK Q.8B ONLY IF "LESS THAN 8 YEARS" IN Q. L OR, - IF "DON'T KNOW/REFUSED" IN Q.M.)

8B. Have any of the toilets been replaced while you've lived there?

Yes _____ - continue
No _____ - Skip to Q. 10
Don't know_____ - Skip to Q.10
Refused _____ - Skip to Q. 10

9. How many toilets have been replaced during this time period?
Record # _____
Don't know _____
Refused _____

10. Based on the information you've given me, your residence is the type we need for the second part of our study. This involves having a technician visit your home to survey the types of water using devices people have in households like yours. Knowing this will allow your local water agency to estimate how much water households like yours will need in the future. This visit will only take about 30 minutes.

After completing the in-home portion of the survey you will be entered into a lottery. Twelve winners will receive a $500.00 gift certificate from a local department store. Your chances of winning are good since this study only includes 800 households.

Interviewer: If respondent asks how many people are eligible for the lottery: Only 800 households in Orange County will be eligible.

If you agree to this visit, you will receive an official confirmation letter from the Metropolitan Water District of Orange County. This letter will include the name of someone you can talk to at the water district office if you have any questions about the study.

Can I have a technician contact you to schedule a convenient time to visit your home for this part of the study?

Yes ___ - continue

No ___ - See below for appropriate response before terminating.

Must check with landlord/manager ___ (Name:_____; Phone:_____), or

Must check with responsible person in household _____ (Name: _____, Relation: _____)
Name of person who completed phone survey (Name: _____, Relation _____)
(*be sure to record both respondent's name and name of responsible person in household if someone else in household completed phone interview and indicated someone such as the parent as the person to contact for in-home inspection). (Obtain respondent's name here, as opposed to very end of survey, only if we need to contact someone other than respondent to schedule the on-site appointment.)*

Refused ___

Interviewer: If respondent asks why the site visit is necessary:
Because you were selected at random and because we know nothing about the housing unit in which you live. Our ability to estimate future water needs and plan the facilities and programs to reliably meet these demands depends critically on knowing what types of water using devices "households like yours" have.

Interviewer: If respondent objects to the interview, please read the appropriate information from the material here:

- ✓ All our staff carry I.D., so you can verify that they represent the Municipal Water Districts of Orange County and Southern California.
- ✓ We will schedule your appointment for a convenient time when you, or someone you designate, will be home and available, 7 days a week, during the day or in the evenings.
- ✓ All the information gathered in our survey is strictly confidential; no one outside our staff will ever see your individual responses.
- ✓ If you want to verify this research, you may call Ms. Pat Meszaros at the Metropolitan Water District of Orange County—the number is 714/ 593-5025. Also, you will

receive a formal letter from the Water District office.
- ✓ The examiner will need only about 30 minutes to complete the inspection.
- ✓ There is no need to check with landlords or apartment managers, since we don't need their approval and the inspection will not damage any appliances or plumbing, so there is no risk to your home. We will call them first, however, if you are uncertain (enter name and phone number where appropriate).

14. When would it be convenient during the next few days to call you (*or the "responsible" person)* to schedule this appointment?

    Day _____ Time _____ **Interviewer: Day should be at least 4 days from today.**

(IF NAME NOT OBTAINED EARLIER IN Q.10, ASK:)

15. May I please have your name so that the technicians know whom to ask for when they call?

    (Mr__ Ms__) First _____Last _____)

VERIFY WITH RESPONDENT THAT THE ADDRESS THEY GIVE IS THE ADDRESS OF THE DIALED RESIDENCE.

Street Address _____
City _____, CA (city may be imported from earlier part of survey)
Zip Code _____ (zip may be imported from earlier part of survey)

A technician will contact you during the next week to schedule an appointment to come out to your home. You will also receive an official confirmation in the mail directly from the Metropolitan Water District of Orange County. Thank you very much for your time and for agreeing to participate in this important study.

APPENDIX D: ON-SITE INSPECTIONS

Design Issues

On-site plumbing inspections were conducted to obtain accurate information about the prevalence of low-flow showerheads and ultra-low-flush toilets. Actual flow rates and flush volumes were measured for this purpose.[73] Once on-site, however, the relatively low marginal cost of collecting additional information about water-using devices and behaviors made this an appealing opportunity. Accordingly, information was gathered about the existence and frequency of use of dishwashers and laundry-machines. Several questions were also included about the presence of pools, spas, outdoor covers, type of irrigation system, and car-washing behavior. Finally, in an effort to develop preliminary estimates of program freeriders, questions were asked about the reasons for replacing toilets and showerheads, and about participation in utility-funded conservation programs.

Questions about dwelling type and age of structure were also included. Although these data had been collected during the telephone survey—to ensure sampling targets were being met—it was expected that the information collected on-site would be more reliable. This was especially important in the case of dwelling type classification. For this reason, the classifications made by the trained plumbing inspectors were taken as definitive.

Among the 762 inspected dwelling units, dwelling type classifications based on information collected during the telephone screening survey and those assigned by the inspectors differed in 87 instances cases. MWDOC staff personally visited 17 of these 87 sites and verified that in 13 cases the on-site inspectors had made the correct determination. Thus, the dwelling-type classifications are thought to be highly reliable. The final determination of structure age (i.e., construction date) was based upon information obtained from Orange County Assessor records. Accordingly, these dates are also assumed to be highly accurate.

Since telephone numbers were dialed at random, inspection-ready households were widely dispersed throughout Orange County at any given point in time. The field survey team was aware and prepared for this. This was an important element of the on-site inspection effort. Alternative scheduling strategies, such as those employing geographic clustering of units to be inspected, could have reduced survey costs by reducing travel time. But this gain would likely have come at the expense of a less geographically representative sample. At the very least it would have substantially complicated the monitoring and the logistics required to assure geographic representativeness.

[73] See Appendix F for a discussion of the measurement procedures used.

On-Site Inspection Statistics

Each household's status was tracked as it changed from the default "not yet contacted" to "on-site visit completed." The following status codes were used:

- Not yet contacted
- Called – No answer
- Called – Said to call back later
- Called – Left message
- Called – Appointment refused
- Appointment scheduled
- Confirmation letter sent
- On-site visit – Appointment refused
- On-site visit – Nobody home
- On-site visit completed
- Hard copy sent to Metropolitan

At the end of the inspection phase of the study, the final case disposition was as follows:

- 762 completed the on-site inspections
- 259 refused to schedule inspections
- 12 changed their minds when the plumbing inspector arrived at their premises
- 12 missed the appointed time and could not be contacted to reschedule.

Table D1 shows inspection completion rates by dwelling type and sample type. Generally SF households exhibited a somewhat lower completion rate, which helped to partially offset the over-representation of SF dwellings enumerated by the telephone survey sample.

Table D1: On-site recruitment and completions, by sample and dwelling type

Sample	Dwelling Type	No. of Households Recruited During the Telephone Survey	No. of Completed Inspections	Completion Rate
City	SF	125	87	69.6%
City	MF	125	94	75.2%
County	SF	551	385	69.9%
County	MF	244	196	80.3%
Totals		1045	762	72.9%

APPENDIX E: ON-SITE INSPECTION DATA COLLECTION INSTRUMENT

Orange County Residential Water Use Survey: On-Site Data Collection Form

Tracking Information

Tracking number: _____
Sample Stratum: _____
Contact name: _____
Contact telephone number: _____
On-site host's name: _____
Appointment date: _____
Appointment time: _____
Street address: _____
City: _____ Zip code: _____
Appraiser: _____
Arrival time: _____ Departure time: _____
Disposition code:

- ☐completed
- ☐ no show
- ☐ breakoff—not rescheduled
- ☐ breakoff/rescheduled➔Date:_____ Time:_____
- ☐ other: _____

Information to be Verified by the Appraiser

Structure Type (*check one*)

- ☐ single family house detached from any other house
- ☐ single family house attached to one or more houses such as a condominium or townhouse (*the units may share a common wall but do not have a separate dwelling unit above or below them)*
- ☐ multi-family building, 2-4 units
- ☐ multi-family building, 5 or more units
- ☐ mobile home or trailer *(a dwelling unit not permanently attached to a foundation)*.
- ☐ other (*describe*): _____

***VERIFY ADDRESS INFORMATION SHOWN UNDER "TRACKING INFORMATION" ABOVE*:**

- ☐ correct
- ☐ incorrect, enter correct address➔ Street name: _____
  City: _____Zip: _____

Is there a separate residence located on a floor directly above or below the one being inspected?:

☐ yes
☐ no

General Questions

I WOULD LIKE TO START WITH A FEW QUESTIONS. SOME OF THESE YOU MAY HAVE ALREADY ANSWERED, BUT I NEED TO VERIFY THIS INFORMATION.

- Do you own or rent? ☐ own ☐ rent ☐ other:_____
- Does your household receive and pay a water bill? ☐ yes ☐ no ☐ don't know
- In what year was this (house/apartment) built:_____
- How long have you lived here?_____ years _____ months
- How may people normally live in this household on a full-time basis?_____
- How many total bathrooms do you have, including those in any outside areas like garages, pool rooms, guest houses and places like that?_____
- How many total bedrooms do you have?_____
- Do you have a dishwasher? ☐ no ☐ yes➔***How may loads of dishes would you say your household does during an average week?_____***
- Do you have a washing machine available for the exclusive use of your household?
  ☐ yes➔***How many loads of laundry would you say your households does during an average week?_____***
  ☐ no ➔***Is there a washing machine on the property available for use by the units in this complex?*** □ yes □ no

Device Inventory and Flow Measurements

AS THE PERSON WHO SPOKE WITH YOU ON THE TELPEHONE SHOULD HAVE EXPLAINED, WE NEED TO INVENTORY AND MEASURE THE WATER FLOW OF SOME OF THE DEVICES YOU HAVE IN YOUR HOME. THIS INFORMATION WILL HELP LOCAL WATER AGENCIES IN PLANNING TO MEET THE FUTURE WATER NEEDS OF HOUSEHOLDS LIKE YOURS.

I'D LIKE TO START WITH THE BATHROOM MOST OFTEN USED BY MEMBERS OF YOUR HOUSEHOLD. CAN YOU SHOW ME THAT BATHROOM?

First Bathroom

Location: ☐ Master bath ☐ Guest bath ☐ Kid's bath ☐ Hallway bath ☐ Outdoor
☐ Other:_____

Has faucet? ☐no ☐ yes➔GPM:_____ **Leaks?**: ☐no ☐ yes

Has shower? ☐no ☐ yes➔GPM:_____ **Leaks?**: ☐no ☐ yes

Has this showerhead been replaced while you've been living here?
☐ no ☐ yes➔***How may years ago was it replaced?***_____
↓
Why did you replace the showerhead?
☐ broken ☐ remodeling ☐ other:_____
↓ ↓ ↓
Did (you/you or your landlord) get this showerhead free of charge, or at reduced cost, through a utility sponsored showerhead replacement program?
☐ no ☐ yes ☐ don't know

Has toilet? ☐ no ☐ yes➔GPF:_____ **Leaks?**: ☐ no ☐ yes

Has this toilet been replaced while you've been living here?
☐ no ☐ yes➔***How may years ago was it replaced?***_____
↓
Why did you replace the toilet?
☐ broken ☐ remodeling ☐other:_____
↓ ↓ ↓
Did (you/you or your landlord) get this toilet free of charge, or at reduced cost, through a utility-sponsored toilet replacement program?
☐ no ☐ yes ☐ don't know
↓

If this program hadn't existed, would (you/you or your landlord) have replaced the toilet yourself at your own cost?
☐ no ☐ yes ☐ don't know

CAN YOU KNOW SHOW ME THE NEXT MOST FREQUENTLY USED BATHROOM?

Second Bathroom

Location: ☐ Master bath ☐ Guest bath ☐ Kid's bath ☐ Hallway bath ☐ Outdoor
☐ Other:_____

Has faucet? ☐ no ☐ yes➔GPM:_____ **Leaks?**: ☐ no ☐ yes

Has shower? ☐ no ☐ yes➔GPM:_____ **Leaks?**: ☐ no ☐ yes

Has this showerhead been replaced while you've been living here?
☐ no ☐ yes➔***How may years ago was it replaced?***_____
↓
Why did you replace the showerhead?
☐ broken ☐ remodeling ☐other:_____
↓ ↓ ↓
Did (you/you or your landlord) get this showerhead free of charge, or at reduced cost, through a utility sponsored showerhead replacement program?
☐ no ☐ yes ☐ don't know

Has toilet? ☐ no ☐ yes➔GPF:_____ **Leaks?**: ☐ no ☐ yes

Has this toilet been replaced while you've been living here?
☐ no ☐ yes➔***How may years ago was it replaced?***_____
↓
Why did you replace the toilet?
☐broken ☐ remodeling ☐ other:_____
↓ ↓ ↓
Did (you/you or your landlord) get this toilet free of charge, or at reduced cost, through a utility-sponsored toilet replacement program?
☐ no ☐ yes ☐ don't know
↓
If this program hadn't existed, would (you/you or your landlord) have replaced the toilet yourself at your own cost?
☐ no ☐ yes ☐ don't know

CAN I NOW SEE THE NEXT MOST FREQUENTLY USED BATHROOM?

Third Bathroom

Location: ☐ Master bath ☐ Guest bath ☐ Kid's bath ☐ Hallway bath ☐ Outdoor
☐ Other:_____

Has faucet? ☐ no ☐ yes➔GPM:_____ **Leaks?**: ☐ no ☐ yes

Has shower? ☐ no ☐yes➔GPM:_____ **Leaks**?: ☐ no ☐ yes

Has this showerhead been replaced while you've been living here?
☐ no ☐ yes➔How may years ago was it replaced?_____
↓
Why did you replace the showerhead?
☐ broken ☐ remodeling ☐ other:_____
↓ ↓ ↓
Did (you/you or your landlord) get this showerhead free of charge, or at reduced cost, through a utility sponsored showerhead replacement program?
☐ no ☐ yes ☐ don't know

Has toilet? ☐ no ☐ yes➔GPF:_____ **Leaks?**: ☐ no ☐ yes

Has this toilet been replaced while you've been living here?
☐ no ☐ yes➔How may years ago was it replaced?_____
↓
Why did you replace the toilet?
☐ broken ☐ remodeling ☐other:_____
↓ ↓ ↓
Did (you/you or your landlord) get this toilet free of charge, or at reduced cost, through a utility-sponsored toilet replacement program?
☐ no ☐ yes ☐ don't know
↓
If this program hadn't existed, would (you/you or your landlord) have replaced the toilet yourself at your own cost?
☐ no ☐ yes ☐ don't know

AND NOW THE NEXT MOST FREQUENTLY USED BATHROOM?

Fourth Bathroom

Location: ☐ Master bath ☐ Guest bath ☐ Kid's bath ☐ Hallway bath ☐ Outdoor
☐ Other:_____

Has faucet? ☐ no ☐ yes➔GPM:_____ **Leaks?**: ☐ no ☐ yes

Has shower? ☐ no ☐ yes➔GPM:_____ **Leaks?**: ☐ no ☐ yes

Has this showerhead been replaced while you've been living here?
☐ no ☐ yes➔How may years ago was it replaced?_____
↓
Why did you replace the showerhead?
☐ broken ☐ remodeling ☐ other:_____
↓ ↓ ↓
Did (you/you or your landlord) get this showerhead free of charge, or at reduced cost, through a utility sponsored showerhead replacement program?
☐ no ☐ yes ☐ don't know

Has toilet? ☐ no ☐ yes➔GPF:_____ **Leaks?**: ☐ no ☐ yes

Has this toilet been replaced while you've been living here?
☐ no ☐ yes➔How may years ago was it replaced?_____
↓
Why did you replace the toilet?
☐ broken ☐ remodeling ☐ other:_____
↓ ↓ ↓
Did (you/you or your landlord) get this toilet free of charge, or at reduced cost, through a utility-sponsored toilet replacement program?
☐ no ☐ yes ☐ don't know
↓
If this program hadn't existed, would (you/you or your landlord) have replaced the toilet yourself at your own cost?
☐ no ☐ yes ☐ don't know

IF UNIT HAS MORE THAN FOUR BATHROOMS, ASK TO SEE EACH REMAINING BATHROOM AND RECORD INFORMATION ON "ADDITIONAL BATHROOMS" FORM.

Wrap-UP

WE'RE ALMOST DONE. I JUST HAVE A FEW MORE QUESTIONS.

- Do you usually take your cars to the car wash, or do you usually wash them yourself?
  - ☐ take them to the car wash
  - ☐ wash myself➔***Considering all of the cars your household has, how may total car washes would you say these cars receive during a typical month?***_____ (total number of car washes)

- Do you have a pool?
  - ☐ yes➔***Does it have a pool cover?*** □ no □ yes
  - ☐ no

  - Do you have a spa?
  - ☐ yes➔***Does it have a cover?*** □ no □ yes
  - ☐ no

Now I'd like to take a look at your outdoor irrigation system. Can you please show me the equipment used to water your yard? (*check all that apply*)

- ☐ no irrigated area
- ☐ buried sprinkler system, manual control
- ☐ buried sprinkler system, automatic control
- ☐ drip, manual control
- ☐ drip, automatic control
- ☐ hose end, manual control

FINALLY, FOR STATISTICAL PURPOSES, CAN YOU PLEASE TAKE A LOOK AT THIS CARD (SHOW INCOME CARD) AND TELL ME THE LETTER OF THE INCOME RANGE THAT INCLUDES THE TOTAL INCOME EARNED BY ALL MEMBERS OF THIS HOUSEHOLD DURING 1999?

- ☐ Income group: _____
- ☐ Refused
- ☐ Don't know

THANK YOU

THANK HOMEOWNER FOR THEIR TIME.

REPEAT THAT THE INFORMATION YOU'VE COLLECTED WILL HELP LOCAL WATER UTILITIES PLAN TO MEET THE FUTURE WATER REQUIREMENTS OF OTHER HOUSEHOLD LIKE THEIRS IN ORANGE COUNTY.

ASK FOR THEIR APPOINTMENT CONFIRMATION LETTER SO YOU CAN BE SURE THEIR HOUSEHOLD IS ENTERED IN THE LOTTERY DRAWING. IF THEY DO NOT HAVE THEIR ORIGINAL LETTER, SIGN AND DATE YOUR COPY. TELL THEM THAT THEIR TELEPHONE NUMBER WILL BE USED AS THEIR ENTRY NUMBER IN THE LOTTERY DRAWING.

Appendix F: On-Site Inspection and Measurement Procedures for Toilets and Showerheads

The procedures used to measure and classify the water flow rates and volumes used by inspected showerheads and toilets and to identify leaks followed those prescribed in Section F of Metropolitan's *Single Family Residential Water Use Survey Program: Surveyor Training Manual and Reference Guide* (Spring 1997). Relevant sections are excerpted below.

TOILET FLUSH VOLUME

1) Determine Size of Toilet Gallons Per Flush (fgp)

The age of the toilet will give the first clue to the flush volume. Toilets made prior to 1980 are typically 5 or 7 gallons per flush (older models have even greater flush volumes). Beginning about 1980, the first 3.5 gallon per flush low-volume toilets were manufactured. These toilets are the same design as the 5 or 7 gallon models, but use a smaller tank. Effective January 1, 1983, the State Plumbing Code required that all new construction include 3.5 gpf toilets. Beginning about 1988 the 1.6 gallon per flush Ultra-Low-Flush toilets (ULF) first appeared. A revision to the State Plumbing Code, effective January 1, 1992, required that all new construction include 1.6 gpf ULF toilets.

ULF toilets have an entirely different and distinct design and can be recognized by the following features:

- The outline of the curving discharge channel inside the base of the toilet is normally visible on the outside of the base.
- The watermark (the surface area of the water in the bowl) is smaller than the water mark of older, larger flush volume toilets.
- The flushing time is much shorter, and the flushing action lacks the large, swirling motion of the older models.
- There may be a float attached to the flapper valve chain.

When all else fails, check for a 1.6 marking either inside the tank cover or elsewhere on the porcelain.

For non-ULF toilets, a date stamp on the underside of the toilet lid or elsewhere on the porcelain should give a clue to the flush volume of the toilet. The following table should be useful:

Determining Toilet Flush Volume

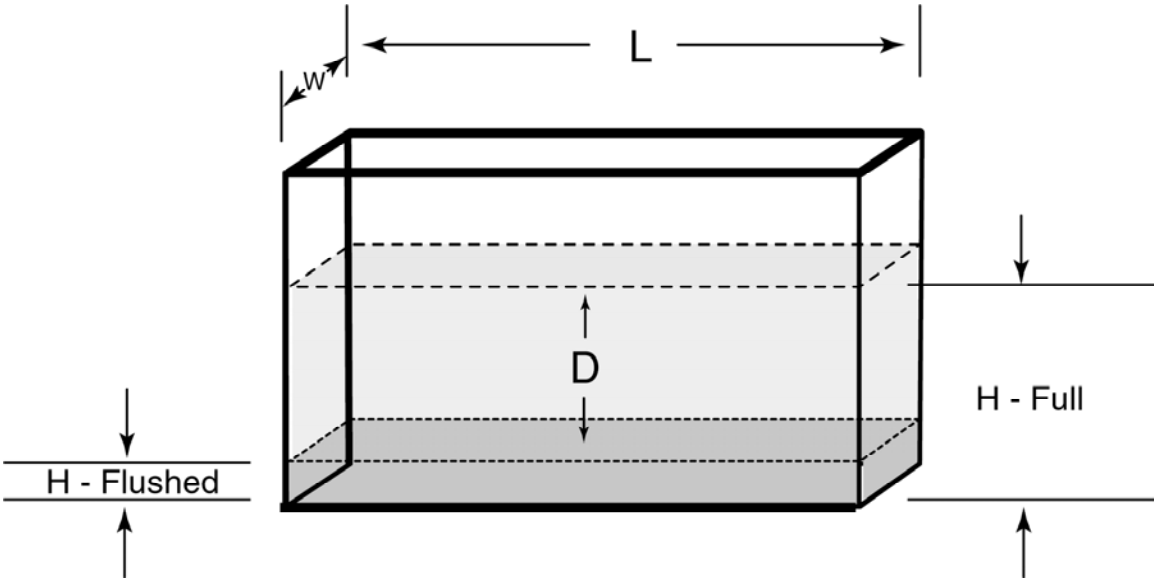
Year of Construction	Anticipated* Size of Toilet (gpf)
Pre 1980	5 – 7
1980-1982	3.5 and larger
1983-1987	3.5
1988-1991	1.6 and 3.5
1992+	1.6

*Actual phase-out of older models may be longer than indicated. Actual toilet currently installed may be newer if home has been retrofitted.

For those situations where it is not clear what the flush volume is, the following measurement technique should identify the flush volume for toilets with tanks that are cubic in shape (i.e., not with sloping sides or sculptured shapes). After removing the lid from the tank:

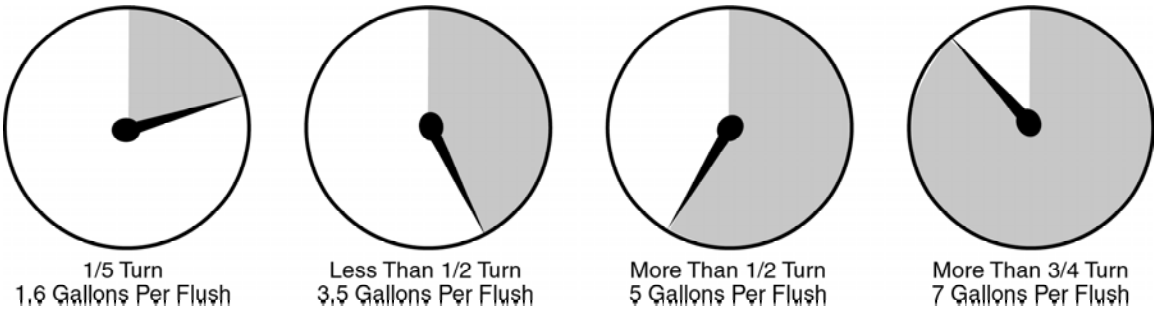
1. measure the dimensions of the tank at its top (L and W)
2. measure height of water in tank when full (H_{full})
3. measure height of water in tank when flushed ($H_{flushed}$)
4. subtract $H_{flushed}$ from H_{full} to get depth of water flushed, D
5. multiply L x W x D
6. divide answer to (5) by 231
7. add 0.5 to answer to (6)

The result is the gallons per flush. ***See the diagram on the following page***.



The following is another more precise method for determining flush volume. Mark the full level on the overflow tube with a grease pencil. Flush the toilet. At the end of the flush, just as the flapper valve returns to the valve seat, prop up the float to prevent the ballcock from releasing water into the tank. Refill the tank manually with a measurement pitcher and record how much water is required to fill the tank to the level marked on the overflow tube. This is the volume of the tank. Add one-half gallon to this volume for the water required to fill the bowl. The total is the gallons per flush.

A final method for determining a toilet's flush volume is to note the amount of water recorded by the water meter when the toilet is flushed. The following is a sample correlation between the meter reading for the one cubic foot dial and the flush volume:



Reprinted, with permission, from the City of San Diego's "How Much Water Does Your Toilet Use?"

Although this method seems simple, coordinating the flush while standing at the meter is sometimes a hassle. Also, the accuracy of the method requires that no other water use is occurring on the site while the toilet is being flushed.

2) Record flush volume

Record one of the following flush volumes on the *Survey Data Collection Form*—Existing GPF: 1.6, 3.5, 5.0, 7.0

Note that the measured flush volumes (regardless of the measuring technique used) will not always correspond precisely to the options above. The actual volume may fall on either side of these values. Choose that volume that corresponds most closely with the measured volume.

SHOWERHEAD REVIEW

A showerhead may use 1.75 to 10+ gallons of water per minute (gpm). It also uses energy to heat the water. Water-conserving showerheads will be installed during the survey to replace all showerheads using more 3.0 gpm.

1) Detect, measure and record leaks

Look for signs of water damage behind the shower wall.

Observe leaks at shower arm, diverter valve, or between shower arm and showerhead. Catch any distinct drips in the calibrated drip gauge for 5 seconds. The calibrations on the side of the gauge convert the amount of water collected to the amount that would drip in a day and in a year. Point out to the customer the size of the leak. Very few people realize the amount of water a leak represents when viewed for a long period of time. The year-long interval specified by the gauge quickly and emphatically demonstrates just how large a seemingly small leak can be.

Record on the Survey Data Collection Form:

- Showerhead leak detected? y/n
- Bathtub diverter leak detected? y/n

2) Measure Flow Rate

Hold flow rate bag or calibrated pitcher over the showerhead. Fully turn on both the hot and cold water faucets, or, if the water are regulated through a single faucet, place the faucet in the mid-position, and turn it on to full force.

Let the water run for five seconds. Use a wristwatch with a second hand or a stop watch to measure the time, or count, “One-thousand-one, one-thousand-two, one thousand-three, one-thousand-four, one-thousand-five”.

The flow rate bag or calibrated pitcher should have gallons per minute marked on its side. (The advantage of a rigid pitcher with a handle is that you can quickly place it under the showerhead with one hand, while observing your watch on the wrist of the other hand.)

After noting the quantity of water, but before pouring it down the drain, ask the customer if there is a plant that needs watering. (Be careful, however, not to make this offer if the house is equipped with a water softener. The salts in the softened water could be harmful to a plant.) This request focuses the customer's attention on the need to save even small quantities of water. If the customer has no such need, pour the water into the sink.

Also, check that the water is completely shut off.

Note: Even though the existing showerhead may be a low-flow model, the flow rate test should be performed to check if the flow restricter has been removed, thereby creating a high-flow showerhead. The book can't be judged by the cover. If the low-flow model has been converted to a high flow model, ask the customer if a replacement showerhead with true low-flow is wanted.

Testing the flow rate can be a wet process, so be sure to carefully wipe the area dry when you're done.

3) Record Flow Rate

Under "Showerheads" on the Survey Data Collection Form in the column designated by you for that bathroom, record the flow rate on the line labeled "Existing gpm(measure)".

Appendix G: Implementation Issues and Solutions

Implementing the *Orange County Saturation Study* survey presented a number of challenges. First, two different firms handled the telephone survey and on-site inspections. This created the need to efficiently coordinate data flows among these and other study parties, and to centrally monitor and adjust, as necessary, the evolving city and county samples. The database application developed to meet these and other needs is discussed in the next section.

Second, there were concerns about how customers would react when asked to grant strangers access to their homes. To assuage these concerns, a number of procedures were developed. The recruitment script used during the telephone screening survey emphasized that the study was being conducted by two public utilities and that the information gathered during the site visit would be used in planning for the region's future water requirements. Recalcitrant participants were given the name and telephone number of a person to contact at the Municipal Water District of Orange County (MWDOC) if they wanted to verify the authenticity of the study. A cumulative list containing the names of all people contacted by the telephone survey company was faxed to all retail water agencies in Orange County in case any of these individual contacted their local water supplier to ask about the study. Once on-site inspection visits were scheduled, confirmation letters were mailed telling participants that the inspector would carry identification and have a duplicate copy of the letter sent to the participant. The confirmation letter also included the name and telephone number of a contact person at MWDOC who could answer questions and verify the legitimacy of the study and the name and telephone number of the plumbing contractor responsible for conducting the on-site inspections. Finally, prospective participants were reminded that, after completing the on-site inspection, they would be eligible for the prize lottery that would award $500 gift certificates to 12 randomly drawn participating households.[74]

Third, as the county and city samples of inspected dwellings evolved, it was necessary to regularly monitor both to ensure sample representativeness and progress in attaining the original completion targets. It was expected that some households who had initially agreed to participate in the on-site inspection phase of the study would later change their minds, meaning that both the representativeness of the telephone enumeration survey and that of the final sample both needed to be monitored. The consistency of dwelling type classifications based on information provided by the telephone survey respondent and the classifications assigned by field inspectors also needed to be carefully monitored. While the inspector's classification was taken as definitive in the analysis phase of the study, sampling adjustments to the telephone enumeration survey often had to rely on apparent discrepancies between the expected mix of single- and multi-family dwellings, developed from Department of Finance housing statistics, and dwelling type classifications based on respondent-reported information.

[74] A copy of the appointment confirmation letter is included at the end of this appendix.

Fourth, since telephone numbers were dialed at random, inspection-ready households were randomly dispersed throughout Orange County at any given point in time. The field survey team anticipated and prepared for this. This turned out to be a sound decision. Alternative strategies involving geographic clustering would have increased the complexity of the study design and increased the difficulties of overall study monitoring and management.

Several situations arose during the data collection process that required remedial action.

The original county sampling plan did not provide for benchmarking the sample in terms of dwelling type due to budget constraints and the expectation that the RDD sampling process would yield a proper share of each type. Halfway through the telephone survey, however, it became apparent that a disproportionate number of single-family dwellings were being recruited largely, it is surmised, because of the greater refusal rate of renter-occupied dwellings—a group comprising the major share of those occupying multi-family dwellings. Fortunately, data collected halfway through the study also suggested that eligible households were agreeing to on-site inspections more frequently than had been assumed during the planning stage (37% versus 15%), which freed up resources that could be used to begin screening for dwelling type and making the necessary adjustments to assure proper proportionate representation of multi-family units in the telephone enumeration survey.[75]

While dwelling type was eventually incorporated as a benchmarking variable for the county sample, date of construction was not. This was partly due to budget constraints, partly because the reliability of the information reported by the telephone survey respondent was expected to be poor, and largely due to the belief that the RDD sample would automatically yield a representative sample of the Orange County housing stock. Once construction date information obtained from the Orange County Assessor's records was combined with site inspector-assigned dwelling type classifications, however, it was discovered that the sample of inspected households included no post-1991 multi-family units. A review of the telephone survey records indicated that these types of units were, in fact, enumerated during the telephone screening survey. But for reasons still unknown, they were not inspected. As discussed in the body of this report, while the systematic exclusion of post-1991 units merits appropriate qualifications regarding the generalizability of study findings, the practical effect is relatively trivial for purposes of the current study. But the outcome underscores the importance of monitoring the yield of sampled

[75] Information about dwelling type and age of structure initially obtained during the telephone survey was later replaced with more reliable information. The on-site determination of dwelling type was used in place of those originally assigned based on information gathered during the telephone survey information. And information from the Orange County Assessor's records was used to determined the age of each inspected structure. To ensure that the on-site determination of dwelling type itself was reliable, MWDOC staff personally visited 17 of the 87 sites where on-site and telephone survey information were inconsistent. They found the on-site information correct in 13 out of the 17 instances. Inspectors were then given remedial instructions on the proper classification of dwelling types. Ultimately, dwelling type codes were discrepant in only 87 cases out of the 762 completed inspections. In all cases the on-site determination was used. For these reasons, the dwelling-type classifications are assumed to be highly reliable.

households in terms of key study variables at both the telephone enumeration and the on-site inspection stages of the study.

Sampling targets were originally established to limit the number of identified study-eligible and willing participants to 100 per week. This was done to make the on-site inspection workload manageable. Immediately after start-up of the field inspection phase, however, the plumbing contractor reported difficulties contacting prospective study participants to schedule the on-site inspection. This problem persisted, despite the fact that calls made to schedule the on-site inspections were made on the day and as close as possible to the time at which the telephone survey respondent has indicated would be the "best" time to reach them to schedule the visit. Because the rate at which appointments were being scheduled and completed was substantially less than had been assumed during the planning stages of study, the plumbing contractor's inspectors lacked sufficient work. Attempting to remedy this situation, the weekly telephone recruitment quota was eliminated in an effort to increase the pool of prospective study households available to be contacted. When this remedy proved inadequate, a second scheduler was added. While the combination of these two adjustments eventually produced enough weekly appointments to keep the inspectors busy, scheduling the on-site appointments was an unexpectedly difficult task and an ongoing source of concern.[76]

Coordinating and Monitoring Data Flows

As the preceding discussion suggests, most of the identified implementation issues involved the need for timely and accurate information flows among the various parties involved in the study. Metropolitan developed a database application in Microsoft® Access for this purpose. Before discussing this application in greater detail, however, it is useful to have an appreciation of the sheer magnitude and complexity of the entire data collection effort.

Figure G1 presents a schematic diagram of the primary information categories and processes, their source of origin, and the way they flowed among study participants. A master legend for interpreting the abbreviations and symbols used in the diagram is given on the following page.

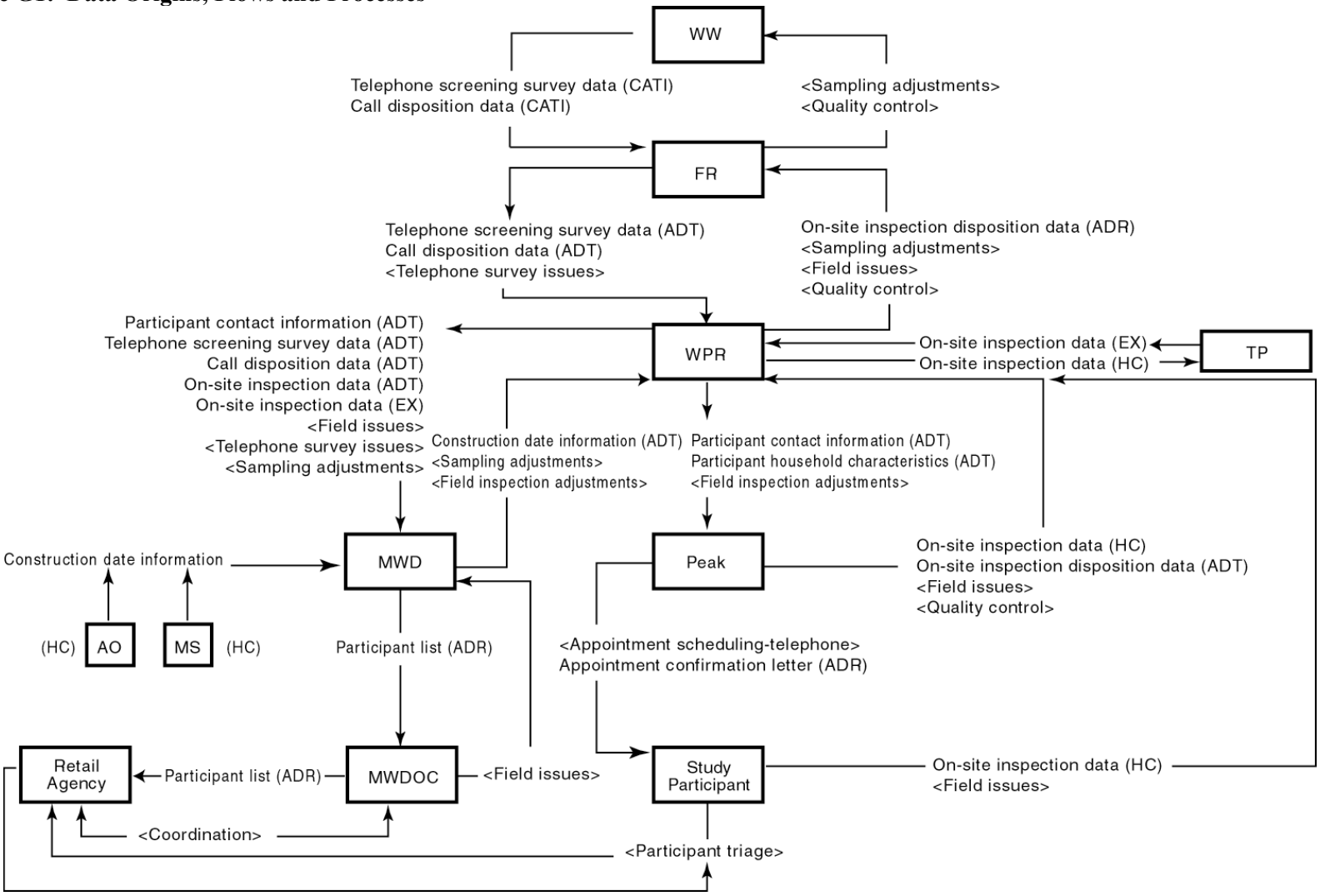
[76] Consideration was given during the planning stage to having the telephone interviewer schedule the on-site inspection appointments. This possibility was soon rejected, however, since it would have instrumentally complicated the logistics involved in coordinating the field inspections. While future studies may want to reconsider the option, the more efficient solution would appear to be increasing the intensity of efforts made to schedule the on-site inspection and allowing for a higher rate of non-contacts and refusals than was assumed by the current study.

Table G1: Legend for Figure G1

Firms, Companies, Organizations and Kindred Entities		Others
AO: Orange County Assessor's Office	TP: TabsPlus (keypunching)	ADR: Access data report
FR: Farrand Research	WPR: Western Policy Research	ADT: Access data table
MS: Metroscan® parcel database	WW: Western WATTS Call Center	EX: Excel table
MWD: Metropolitan Water Disctrict of SC		HC: Hardcopy
MWDOC: Municipal Water District of OC		CATI: Computer-aided telephone interview
PK: Peak International		<…>: process information

Continued on the following page.

Figure G1: Data Origins, Flows and Processes



Database modules were installed at three locations—the project manager's office, the plumbing contractor's office and at MWDOC. Each module had the functionality required by each of these parties. The application had the ability to generate, transmit via email, and import electronic data files and reports. In addition to the ability to generate on-site inspection confirmation letters and the contact lists sent to all retail utilities in Orange County, the database was used to record and report the status of each recruited study household. Additional details about the capabilities and uses of the database application are discussed below.

Project manager. Twice a week, a cumulative file of all households with completed telephone surveys, including those who subsequently refused to participate in the on-site phase of the study, was sent to the project manager. The project manager then imported this file into the database tool and e-mailed an extract to the field survey team. He could also generate reports to use in monitoring the progress and quality of the evolving sample. The extract sent to the field survey team included only the data fields required by them such as the household contact's name, telephone name, address, initial dwelling type classification and the best day and time to contact the household to schedule the visit.

Field survey team. The field survey team imported this cumulative extract into their version of the database tool. The tool was designed to maintain an original copy of the received data, updated every time a new file was imported, and another copy linked by an identifier to which the field survey team could make changes. Data in the latter file was accessible through a data entry window. So, for example, if field survey staff corrected name and street spellings while scheduling appointments, only the second file would be affected. The data entry window also included boxes for entering appointment dates and times as these became available, and dwelling type as determined during the inspections. Finally, each household's status could be recorded as it changed from the default "not yet contacted" to "on-site visit completed." The following status codes were used:

- Not yet contacted
- Called – No answer
- Called – Said to call back later
- Called – Left message
- Called – Appointment refused
- Appointment scheduled
- Confirmation letter sent
- On-site visit – Appointment refused
- On-site visit – Nobody home
- On-site visit completed
- Completed hard copy sent to Metropolitan

As information about newly recruited households arrived, the field survey team would first plot each location on a map. This step was predicated on receiving correctly spelled street addresses,

a point that had to be emphasized to the telephone survey interviewers several times during the early stages of the survey. Once appointment dates and times were finalized and entered into the tool, the click of a button generated appropriately addressed confirmation letters, mailing labels, as well as on-site inspection instruments, the former to be mailed off, the latter for delivery to the plumbing inspectors. The form letter generating function kept a record of processed households so that letters, labels and surveys would not be duplicated every time this function was executed. Automation of these functions eliminated transcription errors. As completed on-site inspection reports were returned to the field head office, inspector-determined dwelling type was added to the database. The rest of the information was key punched much later by a professional data entry firm. The field survey team performed all these tasks on a daily basis.

Twice a week, the field survey team used the database tool to e-mail an extract to the project manager. This extract contained information about the status of each household, appointment dates and times for those scheduled, and dwelling type for those with completed inspections. The project manager used this information to generate an activity status report, including an evaluation of sample representativeness and inspection non-response rates. The project manager also simultaneously e-mailed another extract to MWDOC.

Municipal Water District of Orange County. MWDOC would import this extract into the database tool and then generate either a hardcopy or softcopy list of all inspection-ready households in Orange County recruited to date. For the subset with inspections already scheduled, the list would also show appointment dates and times. This list was forwarded by MWDOC twice a week to all its sub-agencies.

Municipal Water District of Orange County
10500 Ellis Avenue
Fountain Valley, CA 92728

Metropolitan Water District of Southern California
700 North Alameda Street
Los Angeles CA, 90012-2944

«NAME»
«ADDRESS»
«CITY», CA «ZIP»

Dear «NAME»,

On behalf of the Municipal Water District of Orange County and the Metropolitan Water District of Southern California, I want to thank you for participating in the Orange County Residential Water Use Study. I also want to assure you that all information you provide will be kept strictly confidential and will be used only for water planning purposes.

As previously arranged, a technician from Peak International will visit your home on «Appoint_date» at «Appoint_time». The technician will have an identification badge and a copy of this letter. During the visit, the technician will examine and measure the water flow of some of the water using devices in your home. This information will help us in planning how to best meet the water needs of Orange County households like yours during the next 20 years. The technician will also ask a few questions intended to help us understand how households like yours use water. It will save time if you think about your answers to these questions, shown in the box below, before the technician's visit.

How many people live in your home on a full-time basis? _____
How many loads of laundry does your household do during a typical week? _____
If you have a dishwasher, how many loads of dishes do you do during a typical week? _____
If you usually wash your own cars rather than take them to a car wash, how many total car washes would you say you do, including all of your cars, during a typical month? _____

To express our appreciation for your taking the time to participate in this important study, your household will be automatically entered into a lottery after completing the site visit. Twelve of the 800 households participating in the study will win a $500.00 gift certificate for a local appliance or department store. At the conclusion of the site visit, the inspector from Peak International will sign and date this letter, indicating your eligibility for the lottery. Prizewinners will be notified by mail by January 2001.

If you need to reschedule the site visit, please call Peak International (XXX-XXX-XXXX) at least 24 hours before your scheduled appointment. If you have any other questions, please call Ms. XXXXXXXXX (XXX-XXX-XXXX) at the Municipal Water District of Orange County.

Thank you again for participating in this important study

Joe Berg, Water Use Efficiency Programs Coordinator
Municipal Water District of Orange County

Appendix H: Geographic Coverage and Study Participants

As discussed in Section III.B, assessing the quality of the county sample involves two questions. First, how well does the geographic distribution of inspected households agree with the geographic distribution of all Orange County housing units? Second, how well does the group of inspected households match those households who participated in the telephone survey but who declined to participate in the on-site inspection phase of the study?

As Tables H1 and H2 shows, the percentage of inspected households in each city comprising Orange County agrees quite closely with the number that would be expected in a geographically representative sample. Absolute deviations between observed and expected percentages are quite small in most cases. The most notable discrepancies are the relative over-representation of study households in Santa Ana and the corresponding relative under-representation of single-family units in the unincorporated portion of Orange County. The same pattern holds true for multi-family dwellings—Santa Ana is relatively over-represented and dwellings in the unincorporated portion of the county are relatively under-represented. Taken in context, these are minor aberrations in a sample which, taken as a whole, seems quite good in terms of proportionate geographic representation.[77]

Prior to estimating device saturation at the county level, we first examined how well our county sample resembled the housing stock's distribution across cities. We performed this assessment twice; first, by comparing the full county sample to total January 2000 housing stock; and second, by comparing only the pre-1992 sample households to the total January 1992 housing stock (Tables H1 and H2). For example, Table H1 shows that Anaheim accounts for 33 SF inspections, or 8.57% of all inspected SF dwellings. The Department of Finance data indicate that of Orange County's built up SF housing stock as of year 2000, 8.58% were located in Anaheim. Thus, Anaheim's weight in the sample is very close to what one would "expect." Other cities deviate a bit more from these "expected" proportions, but only by small amounts. At a city level, it is difficult to detect any significant, or consistent, pattern of over- or under-representation. At a higher level of aggregation, however, some imbalance becomes apparent. For example, in a perfect world 62.2% and 37.8% of SF inspections should have fallen in northern and southern portions of Orange County, compared to the actual proportions of 66.49% and 33.51% respectively. Among MF dwellings, the imbalance appears a bit greater, northern Orange County accounting for three-fourths of inspected households when ideally this proportion should have been closer to two-thirds. But these imbalances appear minor compared to what is encountered and tolerated in survey work. Table H2 shows how the pre-1992 sample compares to Orange County's pre-1992 housing stock.

Sample weights were developed and used to correct for identified sample imbalances. For example, to bring our SF sample into balance, we down-weighted data from northern Orange

[77] As discussed in Section III, the county sample of inspected households did not include any post-1991 multi-family units. Accordingly, Table H1 assesses the geographic coverage of the multi-family sample in terms of the pre-1992 housing stock.

County by a factor of 0.93 (62.20/66.49) and up-weighted data from southern Orange County by a factor of 1.13 (37.80/33.51). Unincorporated regions were included with south because the vast majority of them are located there. Similarly, to bring the MF sample into balance, we down-weighted data from northern Orange County by a factor of 0.92 (69.61/75.51) and up-weighted data for southern Orange County by a factor of 1.24 (30.39/24.49) (for the MF sample, data from Table H2 are used since the sample only includes pre-1992 households). The ratio of high-to-low weights works out to roughly 1.21 and 1.35 for the SF and MF samples, respectively, which is less than half of our planning estimate (Appendix A).[78] In general, ratios of high-to-low weights as high as 5 are commonly encountered and tolerated in survey work.

Continued on the following page.

[78] Since tests of geographic uniformity have to be based only upon pre-1992 households, weights for both SF and MF samples are derived from Table H2 for those specific analyses.

Table H1: Geographic distribution of the full county sample of single-family units with respect to the January 2000 housing stock

	Single-Family			
City	Number of Inspected Households	Percent of inspected households	Percent of OC's SF housing stock accounted for by city	North OC?
Anaheim	33	8.57%	8.58%	Y
Brea	7	1.82%	1.45%	Y
Buena Park	13	3.38%	2.57%	Y
Costa Mesa	13	3.38%	3.28%	Y
Cypress	11	2.86%	2.02%	Y
Dana Point	8	2.08%	1.62%	N
Fountain Valley	4	1.04%	2.38%	Y
Fullerton	26	6.75%	4.31%	Y
Garden Grove	17	4.42%	5.06%	Y
Huntington Beach	25	6.49%	7.95%	Y
Irvine	11	2.86%	5.39%	N
La Habra	9	2.34%	1.96%	Y
La Palma	2	0.52%	0.64%	Y
Laguna Beach	6	1.56%	1.45%	N
Laguna Hills	9	2.34%	1.32%	N
Laguna Niguel	14	3.64%	2.76%	N
Lake Forest	10	2.60%	2.45%	N
Los Alamitos	2	0.52%	0.39%	Y
Mission Viejo	23	5.97%	4.47%	N
Newport Beach	6	1.56%	3.62%	N
Orange	25	6.49%	4.65%	Y
Placentia	8	2.08%	1.87%	Y
San Clemente	6	1.56%	2.15%	N
San Juan Capistrano	8	2.08%	1.49%	N
Santa Ana	39	10.13%	6.87%	Y
Seal Beach	5	1.30%	0.92%	Y
Stanton	0	0.00%	0.96%	Y
Tustin	11	2.86%	1.79%	N
Villa Park	0	0.00%	0.33%	Y
Westminster	9	2.34%	3.18%	Y
Yorba Linda	8	2.08%	2.84%	Y
Unincorporated	17	4.42%	9.28%	
Totals	38.5	100.0%	100.0%	
North OC	256	66.49%	62.20%	
South OC + Unincorporated	129	33.51%	37.80%	

Table H2: Geographic distribution of the county sample of pre-1992 single- and multi-family units with respect to the January 1992 housing stock

	Single-Family			Multi-Family			
City	Number of Inspected Households	Percent of inspected households	Percent of OC's SF housing stock accounted for by city	Number of inspected households	Percent of inspected households	Percent of OC's MF housing stock accounted for by city	North OC?
Anaheim	30	8.38%	8.81%	34	17.35%	13.97%	Y
Brea	6	1.68%	1.53%	2	1.02%	1.25%	Y
Buena Park	12	3.35%	2.70%	6	3.06%	2.49%	Y
Costa Mesa	13	3.63%	3.49%	16	8.16%	6.28%	Y
Cypress	11	3.07%	2.07%	0	0.00%	0.98%	Y
Dana Point	7	1.96%	1.65%	3	1.53%	1.71%	N
Fountain Valley	4	1.12%	2.55%	4	2.04%	1.05%	Y
Fullerton	26	7.26%	4.55%	13	6.63%	5.39%	Y
Garden Grove	17	4.75%	5.43%	5	2.55%	4.73%	Y
Huntington Beach	25	6.98%	8.18%	16	8.16%	8.26%	Y
Irvine	11	3.07%	5.12%	5	2.55%	4.31%	N
La Habra	9	2.51%	2.06%	7	3.57%	2.16%	Y
La Palma	2	0.56%	0.68%	0	0.00%	0.32%	Y
Laguna Beach	6	1.68%	1.53%	0	0.00%	1.28%	N
Laguna Hills	9	2.51%	1.12%	4	2.04%	0.62%	N
Laguna Niguel	14	3.91%	2.65%	2	1.02%	1.76%	N
Lake Forest	10	2.79%	2.65%	4	2.04%	1.71%	N
Los Alamitos	2	0.56%	0.42%	0	0.00%	0.63%	Y
Mission Viejo	20	5.59%	3.98%	2	1.02%	1.27%	N
Newport Beach	6	1.68%	3.76%	10	5.10%	4.33%	N
Orange	24	6.70%	4.65%	11	5.61%	3.79%	Y
Placentia	7	1.96%	1.74%	1	0.51%	1.18%	Y
San Clemente	6	1.68%	2.07%	5	2.55%	2.29%	N
San Juan Capistrano	6	1.68%	1.44%	2	1.02%	0.47%	N
Santa Ana	38	10.61%	7.42%	26	13.27%	10.18%	Y
Seal Beach	3	0.84%	0.99%	3	1.53%	2.74%	Y
Stanton	0	0.00%	1.02%	1	0.51%	1.54%	Y
Tustin	9	2.51%	1.50%	6	3.06%	3.66%	N
Villa Park	0	0.00%	0.34%	0	0.00%	0.00%	Y
Westminster	9	2.51%	3.33%	3	1.53%	2.12%	Y
Yorba Linda	8	2.23%	2.82%	0	0.00%	0.55%	Y
Unincorporated	8	2.23%	7.75%	5	2.55%	6.98%	
Totals	358	100.0%	100.0%	196	100.0%	100.0%	
North OC	246	68.72%	64.79%	148	75.51%	69.61%	
South OC + Unincorporated	112	31.28%	35.21%	48	24.49%	30.39%	

Probit regression models were used to answer the second question raised at the beginning of this appendix. Specifically, how closely does the sample of inspected households match the subset of households that participated in the telephone survey but declined to participate in the on-site inspection phase of the study?[79] The regression results in Table H3 address this question.

Table H3: Probit regression results

	County Sample		City Sample	
	SF	MF	SF	MF
Independent Variable	Coeff. (Std. Err.)	Coeff. (Std. Err.)	Coeff. (Std. Err.)	Coeff. (Std. Err.)
Number of residents	0.059*	-0.020	0.031	0.064
	(0.025)	(0.036)	(0.052)	(0.056)
Ownership indicator (1=owner; 0=all others)	-0.131	0.012	-0.154	0.224
	(0.116)	(0.141)	(0.221)	(0.248)
Water bill payer indicator (1=yes; 0=no)	-0.073	-0.043	0.011	0.057
	(0.125)	(0.120)	(0.261)	(0.181)
Number of bathrooms	-0.062	-0.010	-0.046	-0.039
	(0.077)	(0.172)	(0.161)	(0.319)
Number of showers	-0.002	-0.224	0.011	-0.065
	(0.084)	(0.204)	(0.184)	(0.332)
Proportion of toilets that are ultra-low-flush	-0.062	0.039	0.318	0.197
	(0.097)	(0.170)	(0.183)	(0.235)
Proportion of showerheads that are low-flow	0.083	-0.096	-0.172	-0.096
	(0.090)	(0.128)	(0.177)	(0.176)
Pre-1992 vintage indicator (1=yes; 0=no)	0.014	0.132	---	---
	(0.156)	(0.327)		
North Orange County (1=yes; 0=no)	-0.139	0.264	---	---
	(0.083)	(0.139)		
Prob. > chi-square	0.158	0.260	0.693	0.803
Number of households used in model	1323	568	339	286
Number of households with complete telephone interviews	1503	632	368	299

*Significant at the 5% level

[79] Ideally, this regression analysis would be extended to compare study participants with households who never participated in the telephone screening survey. This is impossible, however, since the necessary information is obviously not available for telephone survey non-participants. Still, results discussed in this appendix—the proportionate geographic representation of study households and the apparent similarity of those who participated in the telephone screening survey but declined to participate in the on-site inspection with those who did participate in the on-site inspections—provides strong circumstantial evidence in support of the overall representativeness of the sample of study households.

Each column in Table H2 contains the results from a separate regression. The dependent variable for each regression is a binary variable distinguishing households who participated in both the telephone screening survey and the on-site inspection from those who only participated in the telephone screening survey. The independent variables consist of responses given by both groups during the telephone survey.

As can be seen from Table H2, the only statistically significant difference between any of the two groups involves the number of residents in single-family units in the county sample. The effect of this difference is substantially offset by the statistically non-significant chi-square statistic for the regression. This suggests that when the entire set of household characteristics is considered in combination, there is essentially no statistical difference between single-family households who participated in the telephone survey but not the on-site inspections and those single-family households who participated in the full study. The same basic conclusion holds for the other households of central interest in this study.

References

Applied Management and Planning Group. 1994. *Market Saturation of Residential Water Conservation Devices*. A study conducted for the Metropolitan Water District of Southern California. (August 1994). Los Angeles, California.

Berk, Richard A. 1983. "An introduction to sample selection bias in sociological data." *American Sociological Review*. 40:337-48.

California Department of Finance. October 2001. *California Statistical Abstract*. Sacramento, California. This document can be accessed via the Department's web site at www.dof.ca.gov.

California Urban Water Conservation Council. 2000. *Memorandum of Understanding Regarding Urban Water Conservation in California*. (As amended September 21, 2000). Sacramento, California.

Carmines, Edward G. 1980. *Reliability and Validity Assessment*, Beverley Hills, CA: Sage Publications.

The Demand Management Company with Planning and Management Consultants, Ltd. 1994. *Marin Municipal Water District Water Conservation Baseline Study*. April 6, 1994. Corta Madera, CA.

Hollis, Michael E. 1991. *The Public's Social Policy Preferences: Structure and Antecedents.* Unpublished Ph.D. Dissertation. University of California, Los Angeles

Hollis, Michael, Bamezai, Anil and Pekelney, David. 1998. "The Reliability and Validity of Conservation Measures." A paper presented at the American Water Works Association Annual Meeting, June 21-25. Dallas, Texas.

Mayer, Peter.M., DeOreo, William.B., Opitz, Eva M, *et al*. 1999. *Residential Uses of Water*. A study funded by the AWWA Research Foundation. Denver, CO:

Metropolitan Water District of Southern California. Spring 1997. *Single Family Residential Water Use Survey Program: Surveyor Training Manual and Reference Guide, Section F: Interview Survey Step–by–Step Procedure*. Jointly developed by the Metropolitan Water District of Southern California and the Residential Survey Program Advisory Committee (RESPAC). Los Angeles, California.

Muthen, Bengt and Yang, Jin-Wen. 1993. "Selection and predictive validity with latent variable structures." *British Journal of Mathematical and Statistical Psychology*. 46:255-271.

Schuman, Howard and Presser, Stanley. 1981. *Questions and Answers in Attitude Surveys*. New York, NY: Academic Press.